Heriot-Watt University
Research Gateway

Data-driven surrogates for rapid simulation and optimization of WAG injection in fractured carbonate reservoirs

Citation for published version:
Agada, SS, Geiger, S, Elsheikh, AH & Oladyshkin, S 2016, 'Data-driven surrogates for rapid simulation and optimization of WAG injection in fractured carbonate reservoirs', *Petroleum Geoscience*.
https://doi.org/10.1144/petgeo2016-068

Digital Object Identifier (DOI):
10.1144/petgeo2016-068

Link:
Link to publication record in Heriot-Watt Research Portal

Document Version:
Peer reviewed version

Published In:
Petroleum Geoscience

General rights
Copyright for the publications made accessible via Heriot-Watt Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
Heriot-Watt University has made every reasonable effort to ensure that the content in Heriot-Watt Research Portal complies with UK legislation. If you believe that the public display of this file breaches copyright please contact open.access@hw.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 05. May. 2023

Data-driven surrogates for rapid simulation and optimisation of WAG injection in fractured carbonate reservoirs

Simeon Agada[1*], Sebastian Geiger[1], Ahmed Elsheikh[1], Sergey Oladyshkin[2]

[1]Institute of Petroleum Engineering, Heriot-Watt University, Edinburgh, EH14 4AS, United Kingdom.

[2]Department of Stochastic Simulation and Safety Research for Hydrosystems (IWS/SRC SimTech), University of Stuttgart, Pfaffenwaldring 5a, 70569 Stuttgart, Germany.

* Corresponding author.

Now at Imperial College London, United Kingdom (e-mail: s.agada@imperial.ac.uk)

Abstract:

Conventional simulation of fractured carbonate reservoirs is computationally expensive because of the multiscale heterogeneities and fracture-matrix transfer mechanisms that must be taken into account using numerical transfer functions and/or detailed models with a large number of simulation grid cells. The computational requirement increases significantly when multiple simulation runs are required for sensitivity analysis, uncertainty quantification and optimisation. This can be prohibitive, especially for giant carbonate reservoirs. Yet, robust sensitivity analysis, uncertainty quantification and optimisation become increasingly important workflow components as they enable us to analyse, determine and rank the impact of geological and engineering parameters on the economics and sustainability of different Enhanced Oil Recovery (EOR) techniques.

We use experimental design to set up multiple screened simulations of a high-resolution model of a Jurassic Carbonate ramp, which is an analogue for the highly prolific reservoirs of the Arab D formation in Qatar. We consider CO_2 water-alternating-gas (WAG) injection, which has been shown to be a successful EOR method for carbonate reservoirs. The simulations were used as a basis for generating data-driven surrogate models for the rapid simulation and optimisation of hydrocarbon recovery and net gas utilisation. We compare response surfaces from polynomial regression to response surfaces generated with polynomial chaos expansion (PCE). PCE allows for non-linear mapping of parameter uncertainty to the predicted results. In the current work, parameter uncertainties affecting WAG modelling in fractured carbonates are evaluated. These include fracture network properties, fault transmissibility configurations, wettability scenarios, and residual trapping due to hysteresis. Effective fracture permeabilities are computed using discrete fracture networks (DFN) for sparsely distributed regional fractures.

The results enable us to adequately explore the parameter space, quantify and rank the interrelated effect of uncertain model parameters on CO_2 WAG efficiency in fractured carbonate reservoirs. The results highlight the first order impact of the fracture network properties, wettability and hysteresis on hydrocarbon recovery and gas utilisation. Furthermore, surrogate (i.e. proxy) models enable us to calculate quick estimates of the probabilistic uncertainty range and to rapidly optimise hydrocarbon recovery and gas utilisation, while, achieving significant computational speed-up compared with conventional fractured reservoir simulation.

Keywords:

Optimisation, WAG Injection, Data-driven Surrogates, Fractured Carbonate Reservoirs

1. Introduction

Carbonate reservoirs contain a significant proportion of the world's conventional and unconventional hydrocarbon resources, commonly estimated at around 60% of global reserves (Burchette, 2012; Agar and Geiger, 2015). Hydrocarbon recovery in carbonates, however, is typically low, due to multiscale heterogeneities and oil- to mixed-wet rock properties (Manrique et al., 2007; Montaron, 2008; Mohan et al., 2011; Agada et al., 2014). Low recovery factors can further be influenced by complex connected high permeability fracture networks which may establish preferential flow paths in the reservoir (e.g., Bourbiaux et al., 2002; Makel, 2007; Spence et al., 2014). The variability in matrix architecture and fracture network connectivity is the main reason why fractured carbonate reservoirs show a large variety of flow behaviours, leading to significant uncertainties in their evaluation, performance prediction and management (e.g., Cosentino et al., 2001; Makel, 2007; Agada et al., 2016).

To account for multiple geological and engineering uncertainties, a large number of numerical reservoir simulations are typically required to adequately explore the parameter space, investigate parameter relationships and optimise hydrocarbon recovery. Sensitivity analysis, uncertainty quantification and recovery optimisation for fractured carbonate reservoirs, however, are computationally expensive because of the multiscale heterogeneities and fracture-matrix transfer mechanisms that must be taken into account using numerical transfer functions and/or detailed models with a large number of simulation grid cells. This is particularly important for CO_2 WAG injection, a successful EOR method for carbonate reservoirs which combines the benefits of gas injection to reduce the residual oil saturation and water injection to improve mobility control and frontal stability (Christensen et al., 2001; Manrique et al., 2007; Azzolina et al., 2015).

One efficient way of reducing the computational cost is by using data-driven surrogate modelling techniques that construct an approximation (or proxy) of the simulation response based on a limited number of simulation runs (Queipo et al., 2005; Forrester and Keane, 2009;

Gogu et al., 2009; Simpson et al., 2008; Oladyshkin et al., 2011; Gogu and Passieux, 2013; Petvipusit et al., 2014). The modelling process typically involves generating an initial surrogate model with a set of full-physics training simulations. Subsequently, an approximate solution to the objective function is obtained by evaluating the data-driven surrogate. For validation purposes, approximate solutions from the data-driven surrogate are compared to model predictions using full-physics simulation (e.g. black oil or compositional simulation). If the comparison shows a mismatch, the data-driven surrogate is iteratively updated with more training runs and testing points added until the mismatch is eliminated (Koziel and Yang, 2011).

In the context of EOR in fractured carbonate reservoirs, data-driven surrogates may be able to provide good approximations of time consuming numerical simulations. The surrogate models can then help to understand the respective dependencies and correlations of uncertain input parameters and contribute to rapid simulation, optimisation and decision making under uncertainty. Geological parameter uncertainties that affect CO_2 WAG injection include the nature and flow significance of faults and subseismic fractures (Bourbiaux et al., 2002; Casabianca et al., 2007; Ramirez et al., 2009) and the role of wettability and hysteresis when controlling imbibition and drainage in the rock matrix (Larsen and Skauge, 1998; Al-Futaisi and Patzek, 2003; Schmid and Geiger, 2013; Ryazanov et al., 2014). Similarly, engineering parameter uncertainties include WAG design parameters such as the flow rate and location of wells, WAG slug sizes and WAG injection ratios.

The current paper presents results of a synergy between design of experiments, data-driven surrogates and optimisation under uncertainty. The novelty of our work is the synergistic application of the aforementioned approaches to EOR simulation and optimisation for heterogeneous fractured carbonate reservoirs. Although the specific experimental design techniques (i.e. Box-Behnken, Latin Hypercube) and optimisation algorithm (i.e. genetic algorithm) are not new, the application of the experimental design – surrogate workflow to the modelling of fractured carbonate reservoirs has not been previously reported. A brief overview of the state of the art for experimental design, data-driven surrogates from polynomial chaos expansion and optimisation is presented in sections 1.1, 1.2 and 1.3.

1.1 Design of experiments

Design of Experiments (DOE) is commonly used for extensive exploration of parameter spaces (Simpson et al., 2008; Koziel and Yang, 2011). Here, DOE is employed to ensure that data-driven surrogates fully explore the parameter space and provide a robust representation of the full-physics simulation model. DOE aims to maximise the amount of information acquired from a minimum number of simulation runs by optimally allocating samples in the design space (Chen et al., 2006; Montgomery, 2008; Simpson et al., 2008; Myers et al., 2009; Koziel and Yang, 2011). DOE employs different sampling methods to identify a subset of experiments from a larger set according to the number of experimental parameters under investigation.

Deterministic experimental designs such as Box-Behnken, fractional factorial and central composite designs are perfectly orthogonal, explore a large region of the search space and are able to capture model non-linearities (Box et al., 1978; Chen et al., 2006). To select input parameters from random distributions, stochastic samplers such as Latin Hypercube (Helton and Davis, 2003) or nearly orthogonal array (Giunta et al., 2003) are frequently used. Stochastic samplers are also called space filling designs because they are not restricted to sample sizes that are specific multiples of design parameters (Stein, 1987; Giunta et al., 2003; Helton and Davis, 2003).

Here, we use the Box Behnken experimental design to generate surrogate training simulations Box-Behnken is a quadratic experimental design that assures global coverage of the parameter space at acceptable computation cost and takes the interaction of input parameters into account. For validation of the surrogates, we generate surrogate testing simulations using the stochastic Latin Hypercube experimental design which can select input parameters from random distributions and explore the parameter space in a non-rigid way.

1.2 Polynomial chaos expansion

Experimental design techniques coupled with data-driven surrogates have been widely used in hydrocarbon recovery (e.g., Friedmann et al., 2003; Cullick et al., 2006; Panjalizadeh et al., 2014) and CO_2 storage (e.g., Ashraf et al., 2013; Li and Zhang, 2014; Wriedt et al., 2014) applications for uncertainty quantification, risk assessment, optimisation and history

matching. One group of data-driven surrogate modelling techniques that has received increasing attention is polynomial chaos expansion (PCE) (Crestaux et al., 2009; Eldred and Burkardt, 2009; Buzzard, 2012; Oladyshkin et al., 2011; Zhang and Sahinidis, 2012; Ashraf et al., 2013; Elsheikh et al., 2014). PCE methods build a polynomial approximation of the model response using an orthogonal polynomial basis. PCE techniques are efficient and provide a high-order accurate way of including non-linear effects in stochastic analysis (Oladyshkin and Nowak, 2012).

PCE techniques are mainly classified into intrusive and non-intrusive approaches. Intrusive approaches such as the stochastic Galerkin methods (Villadsen and Michelson, 1978; Ghanem and Spanos, 1993; Xiu and Karniadakis, 2003; Matthies and Keese, 2005) require manipulation of the underlying partial differential equations that are solved within the reservoir simulator. Non-intrusive approaches do not require manipulation of the governing equations and use the reservoir simulator as a black box. They are hence more straightforward to apply and involve the evaluation of the coefficients in the chaos expansion using a given number of model simulations (Isukapalli et al., 1998; Li and Zhang, 2007; Blatman and Sudret, 2010; Oladyshkin et al., 2011; Zhang and Sahinidis, 2012; Petvipusit et al., 2014).

In this study, we focus on non-intrusive sparse polynomial chaos expansion (sPCE) and arbitrary polynomial chaos expansion (aPCE) in comparison to polynomial regression (PR). Polynomial regression estimates the coefficients for a second-order polynomial by least squares fitting of the data-driven surrogate model to the training data (Myers et al., 2009). Sparse polynomial chaos (sPCE) is an extension of the generalised polynomial chaos which is based on the Askey Scheme (Askey and Wilson, 1985) of orthogonal polynomials (Xiu and Karniadakis, 2003; Blatman and Sudret, 2010; Elsheikh et al., 2014). Arbitrary polynomial chaos (aPCE) techniques have been shown to minimise the subjectivity of input data distributions by directly using the available information in a data-driven formulation of PCE and employing a global polynomial basis for arbitrary distributions of data (Witteveen et al., 2007; Oladyshkin et al., 2011; Oladyshkin and Nowak, 2012; Ashraf et al. 2013).

1.3 Optimisation

In the presence of multiple uncertainties, finding the most favourable combination of uncertain input parameters to obtain an optimum value of the objective function (e.g. oil recovery, gas utilisation factor) is challenging and commonly requires the application of stochastic optimisation algorithms. Stochastic algorithms including simulated annealing (Dowsland and Thompson, 2012), particle-swarm optimisation (Esmin et al., 2015), neighbourhood algorithm (Subbey et al., 2003), differential evolution (Hajizadeh et al., 2011) and genetic algorithm (Sen et al., 1995; McCall, 2005) have been applied to many reservoir engineering problems. Stochastic algorithms incorporate a random component that allows the search during optimisation to move toward worse solutions occasionally, thereby gaining the ability to seek out the global optimum objective function while escaping from local minima (Abdollahzahdeh et al., 2013).

We use the genetic algorithm, a heuristic search and optimisation technique based on natural evolution through selection (Back and Schwefel, 1993; Gen and Cheng, 2000; Eiben and Smith, 2003; McCall, 2005). The algorithm uses selection, crossover, mutation and recombination of individual reservoir models to obtain a new generation of potentially superior individuals based on ranking with a fitness function (i.e. objective function – see section 3.3). The procedure is repeated to obtain multiple generations until an optimum value of the objective function is reached. The genetic algorithm is robust, flexible and easy to adapt to different engineering problems because it uses the objective function value to determine new search steps and does not require gradient information from the optimisation problem. Hence, the genetic algorithm can be applied to optimisation problems for which traditional algorithms fail because of significant non-linearities or discontinuities in the search space. Several studies provide more details about the genetic algorithm (e.g., Michalewicz, 1996; Mitchell, 1999; Gen and Cheng, 2000) and its application (e.g., Back et al., 2000; McCall, 2005; Costa et al., 2014).

1.4 Objective and workflow

The aim of this study is to generate, analyse and compare non-intrusive data-driven surrogate modelling techniques and illustrate their application to the simulation and optimisation of CO2 WAG injection in fractured carbonate reservoirs where multiple geological (e.g. fracture properties), physical (e.g. trapping of the gas phase) and engineering (e.g. well controls) uncertainties are encountered. We seek to show the benefit of surrogate models for faster sensitivity analysis and optimisation of complex EOR methods in fractured reservoirs by overcoming challenges associated with the high computational cost of conventional simulation. Box-Behnken experimental design is used to set up a wide range of simulations of the high-resolution carbonate reservoir model. Subsequently, the simulations are used to build data-driven surrogates. For validation, additional simulations with random design parameters are set up using the Latin Hypercube experimental design and compared to the response of the data-driven surrogates for the same input parameters. The most accurate surrogate model after validation is then coupled with Monte Carlo methods to generate cumulative distribution functions of oil recovery and gas utilisation. Subsequently, the selected surrogate model is employed for optimisation of the objective function using a genetic algorithm.

A summary of the workflow we have used to construct data-driven surrogates for fractured carbonate reservoirs is presented in figure 1. Input data from multiple sources such as seismic surveys, wireline logs, borehole imaging, petrophysics, core analysis, surface and subsurface analogues is used to build a detailed geological model which is then upscaled to a full-physics finite difference simulation model. Full-physics simulation using the minimum and maximum values of uncertain parameters is used to identify and rank input variables with significant impact (i.e. heavy hitters) on the objective function(s). The heavy hitters are then coupled with DOE techniques to generate surrogate models which are validated before they are employed for rapid simulation, optimisation and uncertainty quantification.

This paper is organized as follows. Section 2 describes the reservoir model, matrix properties, fracture characteristics and fluid properties employed in the full-physics flow simulations used to train and test the surrogate models. The set-up of the data-driven surrogate models is discussed in section 3, including the screening of parameters, experimental design,

surrogate modelling methodology, validation approach and the optimisation algorithm. Section 4 demonstrates the prediction of the objective function(s) with adequately trained surrogate models before describing how goodness of fit measures can be used to validate the surrogates. Subsequently, the surrogates are employed for rapid uncertainty quantification and optimisation. Finally, a discussion of the results and the conclusions are presented in sections 5 and 6, respectively.

2. Reservoir Model Description

2.1 Matrix characterisation and fluid properties

In this study, we use a high-resolution flow simulation model of the Amellago Island Outcrop, a middle Jurassic Carbonate ramp in the High Atlas Mountains of Morocco (Pierre et al., 2010; Amour et al., 2013; Agada et al., 2014). The outcrop can be considered as an analogue for the highly productive carbonate reservoirs of the Arab D formation in Qatar (Al-Saad and Ibrahim, 2005; Al-Emadi et al., 2009). Data from real subsurface reservoirs was used to model porosity and permeability for the facies in the outcrop to ensure a realistic distribution of the reservoir properties, while, the architectural elements of the model were obtained from the outcrop analogue. Many heterogeneous lithologies were preserved in the simulation model including mollusc banks, mud mounds, patch reefs, sub-seismic faults and fractures. Previously, a detailed description of the outcrop geology and static modelling (Agada et al., 2014) and the fracture network modelling (Agada et al., 2016) have been presented.

Due to the large number of simulations required to generate different surrogates, a sector of the Amellago outcrop model consisting of 34 x 35 x 36 grid cells (42,840 cells in total) was used to study CO_2 WAG injection in the heterogeneous reservoir (Fig. 2). Each grid cell has dimensions of 15m x 15m x 3m. An inverted 5-spot well pattern was used with a vertical injection well at the centre of the model and four vertical production wells at the corners. CO_2 WAG injection was simulated using a WAG ratio of 1:1 and eight alternate six-month cycles. The injectors and the producers were set to operate at target liquid rates subject to maximum bottom-hole pressure (BHP) constraints of 41,368 kPa and minimum BHP constraints of 16,547 kPa respectively. The reservoir was assumed to have an initial reservoir pressure of

20,684 kPa and a bubble point pressure of 11,367 kPa. Reference densities for CO_2, oil and water were assumed to be 1.35 kg/m^3, 800 kg/m^3 and 1000 kg/m^3, respectively (Table 1).

To account for rock-fluid interactions during full-physics flow simulations, two-phase relative permeability and capillary pressure curves (i.e. saturation functions) are typically utilised. Here, we use saturation functions similar to those generated by Agada et al. (2016) for end-member wettability scenarios (i.e. water-wet to oil-wet) for carbonate reservoirs. The two-phase saturation functions were generated with Corey (1954) relationships, which for oil/water and gas/oil systems can be described as:

$$k_{rw} = k_{rw,max}\left(\frac{S_w - S_{wi}}{1 - S_{wi} - S_{orw}}\right)^{n_w} \quad (1)$$

$$k_{row} = \left(\frac{1 - S_w - S_{orw}}{1 - S_{wi} - S_{orw}}\right)^{n_{ow}} \quad (2)$$

$$k_{rog} = \left(\frac{1 - S_g - S_{org} - S_{wi}}{1 - S_{gi} - S_{org} - S_{wi}}\right)^{n_{og}} \quad (3)$$

$$k_{rg} = k_{rg,max}\left(\frac{S_g - S_{gi}}{1 - S_{gi} - S_{org} - S_{wi}}\right)^{n_g} \quad (4)$$

$$P_{cow} = P_{cow,max}\left(\frac{S_w - S_{wi}}{1 - S_{wi}}\right)^{-1/\gamma} \quad (5)$$

$$P_{cgo} = P_{cgo,max}\left(\frac{S_o - S_{or}}{1 - S_{or}}\right)^{-1/\gamma} \quad (6)$$

where k_r , S and n denote the relative permeability, fluid saturation and Corey exponent, respectively. Subscripts, w, o and g represent water, oil and gas respectively, while, subscripts i and r denote the initial and residual saturations. γ is the pore size distribution index.

Three-phase saturation functions which are important to account for multiphase flow interactions in the three-phase flow regions generated during WAG injection were computed using the Stone II model (Stone, 1973), while, hysteresis in the relative permeabilities during alternate drainage and imbibition cycles was modelled using the Killough (1976) hysteresis model. For fluid displacement processes where the capillary pressure drop is much less than the drop in viscous pressure at the scale of the grid resolution (such as in this study), capillary

pressure hysteresis effects are negligible and therefore not evaluated. Detailed discussions on the selection and application of three phase saturation functions and hysteresis models for reservoir simulation are not within the scope of this paper.

2.2 Fracture characterisation and discrete fracture network

The unique flow behaviour of fractured carbonate reservoirs is due to the interaction between high-permeability low pore volume fractures and the low-permeability high pore volume matrix. Characterisation of the fracture system is therefore critical to ensure accurate reservoir simulations of fractured carbonate reservoirs which form the basis for accurate surrogates. During the investigation of outcrop analogues, fracture characterisation involves evaluating data from detailed geological observations in the context of well-established conceptual models for the evolution of the fracture network. Conceptual models for the fracture system include but are not limited to pervasive background (or regional) fracture systems, fault related fracture systems and bedding related fracture systems (Makel, 2007; Chesnaux et al., 2007; Agada et al., 2016). Here, we assume that the fractures are part of a pervasive background fracture system with volumetric fracture intensities (P32) that vary from 0.05 m^2/m^3 to 0.2 m^2/m^3. The fracture data is obtained from detailed observations of the Amellago outcrop during extensive field mapping using high-resolution photopanels and LiDAR (Light Detection and Radar).

The fractures are modelled using a discrete fracture network (DFN) approach which is thought to capture the connectivity and scale-dependent heterogeneity of fracture systems (Dershowitz et al., 2000; Bourbiaux et al., 2002; Makel, 2007; Spence et al., 2014). Three intersecting fracture sets are evaluated (Fig. 3). On average, the dip azimuth for each fracture set varies between 95, 135 and 165, while, the dip angle varies between 74, 75 and 76 (Fig. 4). The mean fracture length is 20 m, while, the variation of the fracture length with respect to the mean is defined using an exponential distribution. Fracture apertures with a mean of 0.5 mm are used to estimate fracture permeabilities with the cubic law. Fractures are assumed to be open in all scenarios. Vertical injection and production wells intersect fractures in all cases.

Fracture network flow parameters including equivalent permeability tensors and shape factors were obtained by upscaling the fracture networks to the grid cells of the simulation model (Fig. 5). We have chosen to use the modified Oda (1985) DFN upscaling method that is more computationally efficient than flow-based DFN upscaling and accurate for fracture systems with good connectivity. A dual-porosity dual-permeability formulation (e.g., Kazemi et al., 1992; Bourbiaux et al., 2002) was used to couple fracture-matrix fluid flow due to the significant heterogeneity and hydraulic continuity in the matrix. The exchange of fluids between the fractures and the matrix was modelled using the Gilman and Kazemi (1983) transfer function.

3. Setup of data-driven surrogate models

Data-driven surrogates were generated for two objective functions: the oil recovery factor and net gas utilisation factor (GUF). The oil recovery factor indicates the fraction of oil that is recovered from the reservoir, while, the GUF indicates the net amount of gas that is injected into the reservoir per barrel of oil produced from the reservoir. In general, it is economically desirable to maximise oil recovery and minimise GUF.

The equations used to generate data-driven surrogates with polynomial regression and polynomial chaos expansion are presented below. We assume that second-order polynomials are sufficient to capture the non-linear interactions of the uncertain input parameters in this study. Higher-order polynomials can be employed to incorporate more non-linearity at greater computational expense. The general equation for second-order polynomial regression is given by:

$$f(x) = c_o + \sum_{i_1=1}^{N} c_{i_1} x_{i_1} + \sum_{i_1=1}^{N} c_{i_1 i_1} x_{i_1}^2 + \sum_{i_1=1}^{N} \sum_{i_2=2}^{N} c_{i_1 i_2} x_{i_1} x_{i_2}, \quad (10)$$

where $f(x)$ is the objective function, x_i are the uncertain parameters, c_o is the intercept, c_{i_1} are the coefficients of the linear terms, $c_{i_1 i_1}$ are the coefficients of the quadratic terms; and $c_{i_1 i_2}$ are the coefficients of interaction terms.

The polynomial chaos expansion for a model output Ω is given by:

$$\Omega(x) = \sum_{i=1}^{M} c_i \Psi_i(x), \tag{11}$$

where the coefficients c_i represent the dependence of the model output Ω on the input parameters x. The function Ψ_{i} is a simplified form of the multivariate orthogonal polynomial basis for x. The number of M terms in the expansion depends on the total number of input parameters N and the order d of the expansion, according to equation (12) (Oladyshkin et al., 2011; Hosder, 2012).

$$M = (N + d)!/(N!\, d!) \quad . \tag{12}$$

Subsequently, the unknown coefficients in the expansion (eqn. 2) are evaluated using a non-intrusive least-square collocation method (Moritz, 1978; Chen et al., 2009). For arbitrary polynomial chaos expansion, the data-driven polynomial basis for one random variable (x_j) of degree k is given by:

$$\mathrm{P}_j^{(k)}(x_j) = \sum_{i=0}^{k} p_{i,j}^{(k)}\, x_j^i, \qquad k = \overline{0,d}, \qquad j = \overline{0,N} \tag{13}$$

Here $p_{i,j}^{(k)}$ are the coefficients in $\mathrm{P}_{\mathrm{j}}^{(\mathrm{k})}(x_{\mathrm{j}})$. The coefficients $p_{i,j}^{(k)}$ are constructed in such a way that the polynomials in equation (13) form a basis that is orthogonal in arbitrarily given distributions of data (Oladyshkin et al., 2011). A detailed description of the polynomial basis functions used in sparse polynomial chaos expansion is presented in Elsheikh et al. (2014).

3.1 Parameter screening

Parameter screening is usually the first step in the process of generating surrogate models. Here, full-physics simulation using the minimum and maximum values of uncertain parameters is employed to identify and rank input variables with significant impact (i.e. heavy hitters) on the objective function(s). The heavy hitters are then coupled with experimental design techniques to generate surrogate models. Sensitivity analysis carried out by varying one parameter at a time is a simple and well known procedure for parameter screening. The screening results indicate that the most important uncertainties affecting CO_2 WAG injection

346 in this reservoir include the fracture permeability, matrix wettability (KR), fault
347 transmissibility (FT) and trapped gas saturation (S_{gt}) (Fig. 6).

348 The screening study shows that as uncertain parameters vary between their minimum and
349 maximum values, increasing the fracture permeability typically results in up to a 16% decrease
350 in the oil recovered and the GUF. Conversely, increasing the maximum trapped gas saturation,
351 wettability or fault transmissibility increases the oil recovery (and GUF) by 15%. Only
352 uncertainties that show significant impact on the simulation model response as indicated in
353 figure 6 are considered in the subsequent experimental design and surrogate model set-up.

354

355 **3.2 Experimental design**

356 A Box-Behnken design (Box et al., 1978) was used to vary the uncertain parameters (Table 2).
357 Identical well configurations, flow rates and pressure constraints were maintained to ensure
358 that the variability in simulation outcomes was due to the main uncertain parameters.

359 Fracture permeability multipliers were varied between 0.1 and 10 to account for end-member
360 fracture permeability scenarios. The fault transmissibility was varied between low
361 transmissibility scenarios where the faults were completely sealing (FT = 0) and high
362 transmissibility scenarios where the faults were fully conductive (FT = 1). Relative
363 permeability and capillary pressure curves varied from oil-wet to water-wet corresponding to
364 the low and high end-members respectively. The trapped gas saturation varied from zero (no
365 hysteresis) to a maximum trapped gas saturation of 0.4.

366

367 **3.3 Surrogate modelling and validation**

368 Full-physics reservoir simulations were carried out employing the Box-Behnken experimental
369 design using a training data set of 312 samples. The simulation input variables and the
370 corresponding outputs were used to train polynomial regression (PR), sparse polynomial
371 chaos (sPCE) and arbitrary polynomial chaos (aPCE) algorithms to generate approximations of
372 the simulator output. To test the prediction accuracy of the surrogate models, we evaluated
373 validation simulations using 105 Latin Hypercube samples and compared the response of the

data-driven surrogates to the numerical simulation output. We used the coefficient of determination (R^2), adjusted coefficient of determination (R^2_{adj}) and root mean square error (RMSE) as goodness of fit measures. R^2 indicates how well the data-driven surrogates predict full-physics simulation results. R^2_{adj} is a modified form of the coefficient of determination that accounts for the number of regression coefficients in the surrogate equation. RMSE is the root mean square error of the data-driven surrogate response compared to the full-physics simulation. In general, higher values of R^2, higher values of R^2_{adj} and lower values of RMSE indicate higher surrogate accuracy. Mathematically, R^2, R^2_{adj} and RMSE are given by:

$$R^2 = 1 - \frac{\sum_i^N (y_i - f_i)^2}{\sum_i^N (y_i - \bar{y})^2} \quad (7)$$

$$R^2_{Adj} = 1 - \frac{\sum_i^N (y_i - f_i)^2}{\sum_i^N (y_i - \bar{y})^2} \times \frac{N-1}{N-K} \quad (8)$$

$$RMSE = \sqrt{\frac{\sum_i^N (y_i - f_i)^2}{N}} \quad (9)$$

where y denotes the full-physics simulation result (i.e. oil recovery factor or GUF) used to train the surrogates. $\bar{y}$ is the mean value of N full-physics simulation results evaluated at the end of production. f represents the surrogate predictions corresponding to N simulation cases. K denotes the number of regression parameters utilised in the surrogate model. By incorporating the number of regression parameters, R^2_{adj} provides a conservative estimate of the surrogate accuracy.

3.4 Optimisation with genetic algorithm

The surrogate models were coupled with the genetic algorithm to optimise the oil recovery and GUF based on a modelling framework in which multiple realisations of the geological model are considered while varying operational (i.e. engineering) parameters such as well locations and flow rates to optimise the oil recovery and GUF. Here, we assume multiple realisations of the geological model are obtained when different combinations of the DFN model, saturation functions, residually trapped fractions and fault transmissibility interact

with the matrix, based on the experimental design. Therefore, each combination represents a unique fracture-matrix geological model scenario. Subsequently, the operational parameters of the central injector in the 5-spot well pattern are varied to optimise the oil recovery and GUF across the full range of fracture-matrix geological scenarios. During the optimisation process, the location of the central injector is varied within an area of 120 m^2, while, injection rates are varied up to a maximum of 1987 m^3/day, set to ensure that the well bottom-hole pressures generated during injection are below the formation fracture pressure at all times.

The genetic algorithm optimises an objective function by a process of selection, mutation and recombination as shown in Algorithm 1 (Koziel and Yang, 2011). We used a population size of 50 and a crossover probability of 0.8 to ensure that the algorithm captured a large search space and to avoid being trapped in local minima. Larger population sizes had no effect on the optimisation results. The algorithm was evaluated for 50 generations (i.e. iterations) to obtain optimum results based on a function tolerance of 10^{-6}. The function tolerance defines the minimum difference between new and existing optimal values so that the optimisation iteration is terminated when a predefined function tolerance is reached.

4. Results

4.1 Surrogate training with full-physics simulations

We use black oil simulations in IMEX™ as a basis for generating the data-driven surrogates. The full-physics flow simulations indicate channelling during hydrocarbon displacement in the reservoir which makes CO_2 WAG injection a desirable recovery option because WAG injection can ensure better mobility control and frontal stability to improve contact of injected fluids with unswept zones (Fig. 7a). Buoyant CO_2 migration to the top of the reservoir due to gas-oil density difference is also apparent (Fig. 7b). Furthermore, the full-physics simulations provide the relevant training and testing data sets for generating the proxy models. On average, the computational cost for each black oil simulation run is 8.2 hrs when the simulation is truncated after 1500 days. Considering that simulations were evaluated for 312 Box-Behnken samples and 105 Latin Hypercube samples, truncating each simulation after 1500 days

seemed to be the most feasible way to complete the entire study within a reasonable time frame.

The oil recovery and GUF profiles for the training simulations (Fig. 8a, b) show a range of simulation responses based on various combinations of uncertain input parameters. As expected, the oil recovery increases as alternate cycles of water and gas are injected into the reservoir. The GUF, however, increases initially but begins to decrease as the reservoir becomes gas saturated.

4.2 Oil recovery surrogate prediction

The response surfaces that can be generated from training simulations using the three data-driven surrogate models (PR, sPCE and aPCE) are very similar and the relative error between response surfaces is approximately 0.002. For analysis, we focus on second-order aPCE response surfaces (Fig. 9). We observe from the four response surfaces that the horizontal fracture permeability always has the highest impact on the simulated oil recovery. This clear link between an increase in the fracture connectivity and a decrease in the oil recovery is to be expected because an increased connectivity across the fracture network results in a reduction in the residence time of injected fluids and subsequently a reduction in the effectiveness of oil recovery from the matrix due to gravity drainage and capillary imbibition.

Consequently, the highest overall oil recovery is observed when the fracture permeability is low and the matrix is water-wet and hence imbibition is most effective (Fig. 9c). The lowest overall recovery is observed when both the vertical and horizontal fracture permeabilities are at their highest values (Fig. 9d) indicating that when the fractures are well connected, fracture networks form fluid flow highways that lead to rapid transport of injected fluids thereby resulting in low oil recovery. Increased fault transmissibility (Fig. 9a) allows the injected fluids to access all parts of the reservoir more readily which improves recovery. Similarly, an increase in the maximum trapped gas saturation reduces the overall gas mobility and leads to improved recovery predictions (Fig. 9b). This is because a reduction in the gas mobility increases the stability of the gas-water mobility front, delays gas breakthrough and improves the contact of gas with residual oil, thereby ensuring better microscopic and macroscopic sweep of the reservoir. On average, the computational cost for each surrogate model

evaluation is 13.2 seconds indicating significant reduction in CPU time when compared with the 8.2 hrs CPU time required for a single full-physics simulation. However, consideration must be given to the overhead associated with creating the surrogates. The overhead for creating the surrogates is directly proportional to the number of training and testing simulations that are required to generate robust surrogates. Once the simulations are run, computer codes in MATLAB are applied to the data to generate surrogates within seconds. It is difficult to quantify the time required to write MATLAB codes or analyse the results at each level of modelling complexity as these depend on the experience or expertise of the modeller. For a modeller who fully understands the workflow, a minimum of 7 days simulation using a high performance computer cluster with 20 processors would be required to generate training/testing simulations and generate the surrogate models in this study.

4.3 Gas utilisation factor surrogate prediction

The net gas utilisation factor (GUF) generally increases with increasing horizontal fracture permeability (Fig. 9). This increase is caused by high-permeability fracture networks that allow more gas flow per barrel of oil recovered from the matrix due to the rapid fluid transport in the fractures. We notice that the fault transmissibility has a limited effect on the GUF (Fig. 10a). This is because the fault transmissibility impacts oil and gas migration in the reservoir in the same way: when the fault transmissibility is low, flow of gas and oil across the faults is limited; when the fault transmissibility is high, flow of gas and oil across the faults is enhanced.

The GUF increases with higher values of gas trapping due to hysteresis (Fig. 10b). It is well known that relative permeabilities depend on the saturation path during hydrocarbon displacement cycles (e.g., Larsen and Skauge, 1998). The cycle dependence influences the amount of gas trapped in the subsurface, thereby resulting in higher GUFs as the trapped gas fraction increases. Conversely, the GUF decreases with increasing water-wetness (Fig. 10c). Although the amount of trapped non-wetting gas is higher in a water-wet scenario, the oil recovery is also very high (Fig. 9c). Hence, the GUF, which is a ratio of net gas utilised to oil produced, decreases with increasing water-wetness. The GUF is highest (Fig. 10d) when the

vertical and horizontal fracture permeabilities are high, which indicates rapid gas transport and accumulation at the top of the reservoir when the fracture permeability is very high.

4.4 Surrogate validation: Goodness of fit measures

To validate the surrogate models that were obtained from the training simulation, we compare the predictions of the surrogates with results from full-physics simulations and generate the relevant cross-plots to estimate goodness of fit measures. The coefficient of determination (R^2) for oil recovery obtained from polynomial regression (PR), sparse polynomial chaos (sPCE) and arbitrary polynomial chaos (aPCE) is 0.9635, 0.9768 and 0.9770, respectively (Fig. 11 and Table 3). The R^2 value indicates that all the data-driven surrogates are valid and that the PCE models yield a slightly better approximation of the actual simulation model. The goodness of fit measures for the GUF also show that the PCE models give consistently better predictions of the actual simulation results (Fig. 11 and Table 3). A comparison of the PCE models for both oil recovery and GUF indicates that the aPCE models give marginally better results compared to the sPCE models. However, it is expected that further tuning of the sPCE model may allow us to eradicate the difference between the aPCE and sPCE model. Subsequent relative error analysis, Monte Carlo simulations and model optimisation focus on proxy models from aPCE.

4.5 Surrogate validation: Relative error

Relative error response surfaces (Fig. 12 and 13) show the discrepancy between the response surfaces from PR and aPCE. In comparison to aPCE, PR always over predicts the oil recovery (Fig. 12) and under predicts the GUF (Fig. 13). Analysis of the relative error between the aPCE and PR response surfaces shows that although the overall error is minimal, the difference in the prediction is most evident in the middle of the design space. This is because the deterministic Box-Behnken experimental design used in setting up the training simulations generates samples that more adequately capture the actual model behaviour at the boundaries of the design space but have greater uncertainty at the middle of the design space.

To further investigate the deterministic sampling bias, we generated test simulations using the more random Latin Hypercube experimental design (Fig. 14). We observe that when random samples are added to the design, the mismatch between PR and aPCE prediction has a wider spread in the design space. However, the absolute error from such a random design is greater than the error from the deterministic design.

The final choice of what design method to employ should be a function of how well the surrogate predicts the behaviour of the actual simulation in any given scenario. Furthermore, combining different experimental design techniques, as we have done in this study, could also be a reliable way to account for uncertainties that may propagate from the experimental design techniques used to generate the data-driven surrogates.

4.6 Surrogate based uncertainty quantification and probabilistic assessment

Monte Carlo simulations carried out using the aPCE surrogate and evaluated 65000 times were used to determine the cumulative distribution functions for oil recovery and gas utilisation factor over the range of uncertainty for the input parameters (Fig. 15). The 10^{th}, 50^{th} and 90^{th} (P10, P50 and P90) percentile probabilistic estimate for oil recovery is 0.31, 0.34 and 0.37 respectively for simulation of immiscible CO_2 WAG injection. Also, the P10, P50 and P90 probabilistic estimate is 0.45, 0.53 and 0.60 for GUF.

4.7 Surrogate based optimisation

The aPCE surrogate model coupled with the genetic algorithm was employed to optimise the oil recovery and GUF. Optimisation using the genetic algorithm progresses as a minimisation of the fitness value (i.e. -1 x objective function) with the mean fitness value improving during each generation until the optimum is reached after 50 generations as determined by the predefined function tolerance (Fig. 16).

As discussed in section 3.4, the aPCE surrogate is coupled with the genetic algorithm to optimise the oil recovery and GUF based on a framework where multiple realisations of the geological model are considered while varying operational parameters such as well locations and flow rates (Table 4). It is assumed that multiple realisations of the geological model are

obtained when different combinations of the DFN model, wettability scenario, residually trapped fraction and fault transmissibility interact with the matrix, based on experimental design with each combination representing a unique fracture-matrix scenario. Here, the operational parameters of the central injector in the 5-spot well pattern (Fig. 7) are varied to optimise the oil recovery and GUF across the full range of fracture-matrix geological scenarios. Figure 17 illustrates convergence of the oil recovery (and GUF) to the optimum after 2000 evaluations of the surrogate model based on the genetic algorithm.

When the surrogate-based optimisation results are compared to evaluations of the full-physics model using the optimum input parameters, an absolute error of 0.0048 and 0.0043 is obtained for the oil recovery and GUF respectively. We observe a few random sub-optimal solutions as the algorithm evolves and converges to the optimum due to the random component in the genetic algorithm that allows the search during optimisation to move toward sub-optimal solutions occasionally in order to seek out the global optimum objective (Fig. 17). These random solutions increase our confidence that the algorithm adequately explores the parameter space and obtains a global optimum.

In this study, it was sufficient to optimise a single objective (e.g., oil recovery). Since the oil recovered is inversely proportional to the GUF, maximizing the oil recovery concurrently minimises the GUF which are both desirable outcomes. To study the possibility of optimising many competing objectives, however, multi-objective optimisation is required. Multi-objective optimisation finds a set of optimal solutions in the range between two (or more) optima. The set of optimal solutions, known as the pareto front, should ideally have a good spread (Mohamed et al., 2011; Deb, 2014). The surrogates generated in this study can be utilised for multi-objective optimisation at no additional cost (i.e. no additional simulation runs).

5. Discussion

Reservoir simulation and optimisation of CO_2 WAG injection in fractured carbonate reservoirs is a complex and time-consuming process. By applying surrogate models to approximate full-physics numerical simulations using a limited number of training and testing simulations that cover the parameter space and account for key uncertainties, we can significantly reduce the

overall modelling time. The surrogates can then help to understand the respective dependencies and correlations of uncertain input parameters and contribute to rapid simulation and optimisation under uncertainty.

Response surfaces generated using surrogate models show that fault transmissibility, fracture network properties, matrix wettability, residual trapping due to hysteresis and the fracture network properties are key uncertainties that significantly impact the prediction of oil recovery and gas utilisation for fractured carbonate reservoirs. Furthermore, the interrelated effect of these uncertain parameters is often greater than the impact of one parameter on the model outcome. For example, the interrelated effect of high wettability and low fracture network permeability on oil recovery, is higher than the end-member effect of either of these parameters on oil recovery. Such observations necessitate the application of experimental design techniques that improve evaluation of the parameter space and capture the interactions of major uncertainties. Here, Box-Behnken and Latin Hypercube experimental designs were used to generate a large number of training and testing samples (i.e. full-physics simulations), respectively.

The chosen experimental design is a source of uncertainty in the surrogate modelling workflow which may propagate to the surrogate model prediction because deterministic designs could be biased towards the boundaries of the design, while, random designs may need more training and testing to constrain. By combining deterministic (Box-Behnken) and random (Latin Hypercube) experimental designs to account for the uncertainty from sampling bias, the workflow employed in this study improves the reliability of the surrogate model predictions.

Although, it is considerably faster to evaluate a data-driven surrogate than to run a full simulation case, it is self-evident that such a simple model must be constructed and used with care. The accuracy of the model should be thoroughly validated in order to estimate its prediction capability. Hence, the application of appropriate goodness of fit measures, such as the coefficient of determination (R^2) and the root mean square error (RMSE), is essential to ensure that the surrogate reliably replaces the full simulation model inside and outside of the design space. When the surrogates generated in this study are compared using R^2 and RMSE, surrogate results from polynomial chaos expansion (PCE) – both sparse and arbitrary PCE, consistently give better results than traditional polynomial regression.

609 The work presented in this paper, provides a solid basis for diverse applications of PCE-based
610 surrogates to several aspects of fractured reservoir simulation and optimisation that would
611 benefit from the computationally efficient workflow. First, the PCE-based surrogates can be
612 applied to advanced global sensitivity analysis using Sobol indices (e.g., Buzzard, 2012;
613 Oladyshkin et al., 2012). As discussed in section 3, the PCE-based surrogate output is
614 presented as an orthogonal decomposition through the uncertain input parameters. The
615 orthogonal decomposition can directly be employed through Sobol sensitivity indices (Sobol,
616 1990) to quantify the relative importance of uncertain input parameters on the final
617 prediction. Once the PCE-based surrogate model is generated, the sensitivity indices can be
618 constructed on-the-fly using analytical relations, thereby, providing information on the high
619 order interaction between contributing model parameters (e.g., Oladyshkin et al., 2012).

620 Second, robust optimisation under geological uncertainty (e.g., Mulvey and Vanderbei, 1995;
621 Nghiem et al., 2009; Chen et al., 2012; Petvipusit et al., 2014) can be achieved using the
622 developed surrogates. During robust optimisation, a given objective function is optimised by
623 modifying engineering parameters (e.g., well location and flow rates) for a wide range of
624 geological scenarios, thereby, capturing geological uncertainty in the optimisation process.
625 Typically, robust optimisation progresses by optimising over the average and standard
626 deviation of model results generated with different geological realisations. Because the
627 average response surface obtained during robust optimisation is much smoother than the
628 response surfaces for individual realisations, it can potentially reduce the total number of
629 simulations needed to build surrogates.

630 Third, multi-objective optimisation can be carried out to optimise competing objectives (e.g.,
631 Mohamed et al., 2011; Deb, 2014). For example, the oil recovery and net present value can
632 be maximised while concurrently minimizing the GUF and water cut. When multi-objective
633 optimisation is employed in the framework of geological uncertainty, the objective function
634 will need to reflect the impact of geological uncertainties by using either a mean value or the
635 mean value combined with the standard deviation for each objective. Subsequently, an
636 optimisation algorithm (e.g., the classic genetic algorithm or the more recent Non-dominated
637 Sorting Genetic Algorithm-II) is run on the PCE-based surrogate to obtain a pareto-optimal
638 front representing competing objectives. The accuracy of the optimisation outcome can be

progressively improved by re-training the surrogates along the pareto-optimal front and re-running the optimisation algorithm.

This study seeks to demonstrate how surrogate models for fractured carbonate reservoirs can be coupled with a wide range of reservoir optimisation techniques. Therefore, it should be noted that we do not focus on the details of specific optimisation algorithms. We use the well-known genetic algorithm but more advanced techniques that apply efficient gradient-based or stochastic techniques to field-scale reservoir optimisation have been widely researched (e.g., Dowsland and Thompson, 2012; Isebor et al., 2014; Esmin et al., 2015).

6. Conclusion

The purpose of this study was to generate, analyze and compare non-intrusive data-driven surrogate modelling techniques, and illustrate their application to the simulation and optimisation of CO_2 WAG injection in fractured carbonate reservoirs. The synergistic application of experimental design, data-driven surrogates and genetic algorithms for CO_2 WAG simulation and optimisation represents a notable contribution of this work. We have shown that data-driven surrogates from PCE (arbitrary polynomial chaos expansion, aPCE, and sparse polynomial chaos expansion, sPCE) show a higher degree of accuracy in predicting oil recovery and GUF compared to surrogates from polynomial regression. PCE techniques capture the synergistic effects between low- and high-order polynomial terms and thereby provide higher accuracy. In particular, aPCE most closely approximates the actual simulations when trained and tested.

We demonstrate that data-driven surrogate models significantly reduce the computational cost by completing each model evaluation in 13.2 seconds compared to 8.2 hours for full-physics simulation using the inputs. Hence, we are able to rapidly evaluate the dependency and correlation of uncertain input parameters as they influence the oil recovery and GUF. For example, we find that low fracture permeabilities, more water wetting saturation functions, high residual trapping due to hysteresis and high fault transmissibilities are favourable to achieve higher oil recovery. When the computationally efficient surrogates are coupled with

668 the genetic algorithm, over 2000 model evaluations are rapidly carried out to optimise the oil
669 recovery and show the combination of input variables that are favourable to the optimum
670 recovery scenario.

671

672

673 ## Acknowledgements

674 The authors would like to thank the ExxonMobil $(FC)^2$ Research Alliance for funding this project.
675 Sebastian Geiger is grateful to Foundation CMG for supporting his chair in carbonate reservoir
676 simulation. We acknowledge Computer Modelling Group, Schlumberger and Golder Associates for
677 providing access to IMEX, PETREL and FRACMAN, respectively.

678 ## References

679 Abdollahzadeh, A., Reynolds, A., Christie, M., Corne, D. W., Williams, G. J. J., & Davies, B. J. 2013. Estimation of
680 Distribution Algorithms Applied to History Matching. *SPE Journal*. doi:10.2118/141161-PA
681 Abushaikha, A. S. & Gosselin, O. R. 2008. Matrix-fracture transfer function in dual-medium flow simulation:
682 review, comparison and validation. Paper SPE 113890, presented at SPE Europec Annual Conference and
683 Exhibition, Rome, 9-12 June.
684 Agada, S., Chen, F., Geiger, S. et al. 2014. Numerical simulation of fluid-flow processes in a 3D high-resolution
685 carbonate reservoir analogue. *Petroleum Geoscience*, 20(1), 125-142.
686 Agada, S., Geiger, S. & Doster, F. 2016. Wettability, Hysteresis and Fracture-Matrix Interaction during CO_2 EOR
687 and Storage in Fractured Carbonate Reservoirs. *International Journal of Greenhouse Gas Control,* 46, 57-75.
688 Agar, S. M. & Geiger, S. 2015. Fundamental controls on fluid flow in carbonates: current workflows to emerging
689 technologies. In: Agar, S.M. & Geiger, S. (eds) *Fundamental Controls on Fluid Flow in Carbonates*. Geological
690 Society, London, Special Publications, **406**, http://dx.doi.org/10.1144/SP406.18
691 Al-Emadi A., Jorry S., Chautru, J., Caline, B., Blum, M., Jeddaan, N., Fryer, V., Leandri, P. & Fraisse, C. 2009. 3D
692 Modeling of the Arab Formation (Maydan Mahzam Field, Offshore Qatar): An Integrated Approach.
693 International Petroleum Technology Conference, Doha, Qatar, 7–9 December.
694 Al-Saad, M. & Ibrahim, I. 2005. Facies and palynofacies characteristics of the Upper Jurassic Arab D reservoir in
695 Qatar. *Revue de Paléobiologie, Genève*, **24** (1), 225-241.
696 Amour, F., Mutti, M., Christ, N., et al. 2013. Outcrop analogue for an oolitic carbonate ramp reservoir: A scale-
697 dependent geologic modelling approach based on stratigraphic hierarchy. *AAPG Bulletin*, **97**, 845-871.
698 Ashraf, M., Oladyshkin, S., & Nowak, W. 2013. Geological storage of CO_2: Application, feasibility and efficiency
699 of global sensitivity analysis and risk assessment using the arbitrary polynomial chaos. *International Journal of*
700 *Greenhouse Gas Control*, **19**, 704-719.

Askey, R., & Wilson, J. A. 1985. *Some basic hypergeometric orthogonal polynomials that generalize Jacobi polynomials* (Vol. 319). American Mathematical Society.

Azzolina, N. A., Nakles, D. V., Gorecki, C. D., Peck, W. D., Ayash, S. C., Melzer, L. S., & Chatterjee, S. 2015. CO2 storage associated with CO2 enhanced oil recovery: A statistical analysis of historical operations. *International Journal of Greenhouse Gas Control*, *37*, 384-397.

Bäck, T., & Schwefel, H. P. 1993. An overview of evolutionary algorithms for parameter optimization. *Evolutionary computation*, *1*(1), 1-23.

Bäck, T., Fogel, D. B., & Michalewicz, Z. (Eds.). 2000. *Evolutionary computation 1: Basic algorithms and operators* (Vol. 1). CRC Press.

Blatman, G., & Sudret, B. 2010. An adaptive algorithm to build up sparse polynomial chaos expansions for stochastic finite element analysis.*Probabilistic Engineering Mechanics*, *25*(2), 183-197.

Bourbiaux, B., Basquet, R., Cacas, M. C., Daniel, J. M., & Sarda, S. 2002. An integrated workflow to account for multi-scale fractures in reservoir simulation models: implementation and benefits. SPE 78489, In *Abu Dhabi International Petroleum Exhibition and Conference, 13-16, October.*

Box, G. E., Hunter, W. G., & Hunter, J. S. 1978. Statistics for experimenters. New York: Wiley, 1978.

Burchette, T. 2012. Carbonate rocks and petroleum reservoirs: a geological perspective from industry. Geological Society London, Special Publications, **370.**

Buzzard, G. T. 2012. Global sensitivity analysis using sparse grid interpolation and polynomial chaos. *Reliability Engineering & System Safety*,*107*, 82-89.

Casabianca, D., Jolly, R. J. H., & Pollard, R. 2007. The Machar Oil Field: waterflooding a fractured chalk reservoir. *Geological Society, London, Special Publications*, **270**(1), 171-191.

Christensen, J. R., Stenby, E. H. & Skauge, A. 2001. Review of WAG field experience. *SPE Res. Eval. & Eng.*, **4**, 97-106.

Chen, V. C., Tsui, K. L., Barton, R. R., & Meckesheimer, M. 2006. A review on design, modeling and applications of computer experiments. *IIE transactions*, *38*(4), 273-291.

Chen, J. S., Wang, L., Hu, H. Y., & Chi, S. W. 2009. Subdomain radial basis collocation method for heterogeneous media. *International journal for numerical methods in engineering*, *80*(2), 163-190.

Echeverria Ciaurri, D., Isebor, O. J., & Durlofsky, L. J. 2011. Application of derivative-free methodologies to generally constrained oil production optimisation problems. *International Journal of Mathematical Modelling and Numerical Optimisation*, *2*(2), 134-161.

Cosentino, L., Coury, Y., Daniel, J. M., et al. 2001. Integrated study of a fractured Middle East reservoir with stratiform super-k intervals – Part2: Upscaling and dual media simulation. Paper SPE 68184 presented at the Middle East Oil Show, Bahrain, 17-20 March.

Costa, L. A. N., Maschio, C., & Schiozer, D. J. 2014. Application of artificial neural networks in a history matching process. *Journal of Petroleum Science and Engineering*, *123*, 30-45.

Crestaux, T., Le Maıˆtre, O., & Martinez, J. M. 2009. Polynomial chaos expansion for sensitivity analysis. *Reliability Engineering & System Safety*,*94*(7), 1161-1172.

738 Cullick, A. S., Johnson, W. D., & Shi, G. 2006. Improved and more rapid history matching with a nonlinear proxy
739 and global optimization. In *SPE Annual Technical Conference and Exhibition*. Society of Petroleum Engineers.
740 Deb, K. 2014. Multi-objective optimization. *Search methodologies* (pp. 403-449). Springer US.
741 Dernaika, M. R., Basioni, M. A., Dawoud, A., et al. 2013. Variations in bounding and scanning relative permeability
742 curves with different carbonate rock types. *SPE Res. Eval. & Eng.*, **16**(3), 265-280.
743 Dershowitz, B., Lapointe, P., Eiben, T. & Wei, L. 2000. Integration of discrete feature network methods with
744 conventional simulator approaches. *SPE Res. Eval. & Eng.*, **3**(2), 165 – 170.
745 Dowsland, K. A., & Thompson, J. M. 2012. Simulated annealing. In *Handbook of Natural Computing*. Springer
746 Berlin Heidelberg.
747 Eiben, A. E., & Smith, J. E. 2003. *Introduction to evolutionary computing*. Springer Science & Business Media.
748 Eldred, M. S., & Burkardt, J. 2009. Comparison of non-intrusive polynomial chaos and stochastic collocation
749 methods for uncertainty quantification. *AIAA paper*, **976**, 1-20.
750 Elsheikh, A. H., Hoteit, I., & Wheeler, M. F. 2014. Efficient Bayesian inference of subsurface flow models using
751 nested sampling and sparse polynomial chaos surrogates. *Computer Methods in Applied Mechanics and*
752 *Engineering*, *269*, 515-537.
753 Esmin, A. A., Coelho, R. A., & Matwin, S. 2015. A review on particle swarm optimization algorithm and its variants
754 to clustering high-dimensional data. *Artificial Intelligence Review*, *44*(1), 23-45.
755 Friedmann, F., Chawathe, A., & Larue, D. K. 2003. Assessing uncertainty in channelized reservoirs using
756 experimental designs. *SPE Reservoir Evaluation & Engineering*, *6*(04), 264-274.
757 Forrester, A. I., & Keane, A. J. 2009. Recent advances in surrogate-based optimization. *Progress in Aerospace*
758 *Sciences*, *45*(1), 50-79.
759 Gen, M., & Cheng, R. 2000. *Genetic algorithms and engineering optimization*. John Wiley & Sons.
760 Ghanem, R., & Spanos, P. D. 1993. A stochastic Galerkin expansion for nonlinear random vibration
761 analysis. *Probabilistic Engineering Mechanics*, *8*(3), 255-264.
762 Gilman, J. R., & Kazemi, H. 1983. Improvement of simulation of naturally fractured reservoirs. *SPE Journal*, **23**,
763 695-707.
764 Giunta, A. A., Wojtkiewicz, S. F., & Eldred, M. S. 2003. Overview of modern design of experiments methods for
765 computational simulations. In*Proceedings of the 41st AIAA Aerospace Sciences Meeting and Exhibit, AIAA-*
766 *2003-0649*.
767 Gogu, C., Haftka, R. T., Bapanapalli, S. K., & Sankar, B. V. 2009. Dimensionality reduction approach for response
768 surface approximations: application to thermal design. *AIAA journal*, *47*(7), 1700-1708.
769 Gogu, C., & Passieux, J. C. 2013. Efficient surrogate construction by combining response surface methodology
770 and reduced order modeling.*Structural and Multidisciplinary Optimization*, *47*(6), 821-837.
771 Hajizadeh, Y., Demyanov, V., Mohamed, L., & Christie, M. 2011. Comparison of evolutionary and swarm
772 intelligence methods for history matching and uncertainty quantification in petroleum reservoir models.
773 *Intelligent Computational Optimization in Engineering*, 366, 209-240.
774 Helton, J. C., & Davis, F. J. 2003. Latin hypercube sampling and the propagation of uncertainty in analyses of
775 complex systems. *Reliability Engineering & System Safety*, *81*(1), 23-69.

776 Hosder, S. 2012. Stochastic response surfaces based on non–intrusive polynomial chaos for uncertainty
777 quantification. International Journal of Mathematical Modelling and Numerical Optimisation, 3(1), 117-139.
778 Hui, M. H., & Blunt, M. J. 2000. Effects of wettability on three-phase flow in porous media. *The Journal of Physical*
779 *Chemistry B*, **104**(16), 3833-3845.
780 Isebor, O. J., Durlofsky, L. J., & Ciaurri, D. E. 2014. A derivative-free methodology with local and global search for
781 the constrained joint optimization of well locations and controls. *Computational Geosciences*, *18*(3), 463-482.
782 Isukapalli, S. S., Roy, A., & Georgopoulos, P. G. 1998. Stochastic response surface methods (SRSMs) for
783 uncertainty propagation: application to environmental and biological systems. *Risk analysis*, *18*(3), 351-363.
784 Juanes, R., Spiteri, E. J., Orr, F. M. & Blunt, M. J. 2006. Impact of relative permeability hysteresis on geological
785 CO_2 storage. *Water Resources Research*, **42**(12), 1-13.
786 Kazemi, H., Gilman, J. R., & Elsharkawy, A. M. 1992. Analytical and Numerical Solution of Oil Recovery From
787 Fractured Reservoirs With Empirical Transfer Functions. *SPE Reservoir Engineering*, ***7***(2), 219-227.
788 Killough, J.E., 1976. Reservoir simulation with history-dependent saturation functions. *SPE Journ*al, **16**, 37– 48.
789 Koziel, S., & Yang, X. S. 2011. *Computational optimization, methods and algorithms* (Vol. 356). Germany:
790 Springer.
791 Larsen, J. A. & Skauge, A. 1998. Methodology for numerical simulation with cycle-dependent relative
792 permeabilities. *SPE Journal*, **3**, 163-173.
793 Le Maître, O. P., & Knio, O. M. 2010. *Spectral methods for uncertainty quantification: with applications to*
794 *computational fluid dynamics*. Springer.
795 Li, H., & Zhang, D. 2007. Probabilistic collocation method for flow in porous media: Comparisons with other
796 stochastic methods. *Water Resources Research*, **43**(9).
797 Li, S. & Zhang, Y. 2014. Model complexity in carbon sequestration: a design of experiment and response surface
798 uncertainty analysis. *International Journal of Greenhouse Gas Control*, **22**, 123-138.
799 Makel, 2007. The modelling of fractured reservoirs: constraints and potential for fracture network geometry and
800 hydraulics analysis. Geological Society, London, Special Publications, **292**, 375-403.
801 Manrique, E. J., Muci, V. E. & Gurfinkel, M. E. 2007. EOR field experiences in carbonate reservoirs in the United
802 States. *SPE Reservoir Evaluation & Engineering*, **10**(6), 667-686.
803 Matthies, H. G., & Keese, A. 2005. Galerkin methods for linear and nonlinear elliptic stochastic partial differential
804 equations. *Computer Methods in Applied Mechanics and Engineering*, *194*(12), 1295-1331.
805 McCall, J. 2005. Genetic algorithms for modelling and optimisation. *Journal of Computational and Applied*
806 *Mathematics*, *184*(1), 205-222.
807 Michalewicz, Z. 1996. *Genetic algorithms+ data structures= evolution programs*. Springer Science & Business
808 Media.
809 Mitchell, M. 1999. An introduction to genetic algorithms. Cambridge, Massachusetts: MIT Press, 158 pp.
810 Mohamed, L., Christie, M. A., & Demyanov, V. 2011. History matching and uncertainty quantification:
811 multiobjective particle swarm optimisation approach. Paper SPE 143067 presented at the SPE EUROPEC/EAGE
812 Annual Conference and Exhibition, 23-26 May, Vienna, Austria.

Mohan, K., Gupta, R. & Mohanty, K. K. 2011. Wettability altering secondary oil recovery in carbonate rocks. *Energy and Fuels*, **25**, 3966-3973.

Montaron, B. 2008. Carbonate Evolution. *Oil & Gas Middle East*, August, 26-31.

Montgomery, D. C. 2008. *Design and analysis of experiments*. John Wiley & Sons.

Moritz, H. 1978. Least-squares collocation. *Reviews of Geophysics and Space Physics*, 16(3):421–430.

Myers, R. H., Montgomery, D. C., & Anderson-Cook, C. M. 2009. *Response surface methodology: process and product optimization using designed experiments* (Vol. 705). John Wiley & Sons.

Oda, M. 1985. Permeability tensor for discontinuous rock masses. Geotechnique, **35**(4), 483-495.

Oladyshkin, S., Class, H., Helmig, R., & Nowak, W. 2011. A concept for data-driven uncertainty quantification and its application to carbon dioxide storage in geological formations. *Advances in Water Resources*, **34**(11), 1508-1518.

Oladyshkin, S., de Barros, F. P. J., & Nowak, W. 2012. Global sensitivity analysis: a flexible and efficient framework with an example from stochastic hydrogeology. *Advances in Water Resources*, *37*, 10-22.

Oladyshkin, S., & Nowak, W. 2012. Data-driven uncertainty quantification using the arbitrary polynomial chaos expansion. *Reliability Engineering & System Safety*, *106*, 179-190.

Panjalizadeh, H., Alizadeh, N., & Mashhadi, H. 2014. A workflow for risk analysis and optimization of steam flooding scenario using static and dynamic proxy models. *Journal of Petroleum Science and Engineering*, *121*, 78-86.

Petvipusit, K. R., Elsheikh, A. H., Laforce, T. C., King, P. R., & Blunt, M. J. 2014. Robust optimisation of CO_2 sequestration strategies under geological uncertainty using adaptive sparse grid surrogates. *Computational Geosciences*,*18*(5), 763-778.

Pierre, A., Durlet, C., Razin, P. & Chellai E. H. 2010. Spatial and temporal distribution of ooids along a Jurassic carbonate ramp: Amellago outcrop transect, High-Atlas, Morocco. Geological Society, London, Special Publication, **329**, 65-88.

Quandalle, P. & Sabathier, J. C. 1989. Typical features of a multipurpose reservoir simulator. SPE Reservoir Engineering, November publication.

Queipo, N. V., Haftka, R. T., Shyy, W., Goel, T., Vaidyanathan, R., & Kevin Tucker, P. 2005. Surrogate-based analysis and optimization. *Progress in aerospace sciences*, *41*(1), 1-28.

Ramirez, B., Kazemi, H., Al-Kobaisi, M., Ozkan, E. & Atan, S. 2009. A critical review for proper use of water/oil/gas transfer functions in dual-porosity naturally fractured reservoirs: Part I. *SPE Res. Eval & Eng*, **12**(2), 200-210.

Ryazanov, A. V., Sorbie, K. S., & van Dijke, M. I. J. 2014. Structure of residual oil as a function of wettability using pore-network modelling. *Advances in Water Resources*, 63, 11-21.

Schmid, K. S., & Geiger, S. 2012. Universal scaling of spontaneous imbibition for water-wet systems. *Water Resources Research*, *48*(3).

Sen, M. K., Datta-Gupta, A., Stoffa, P. L., Lake, L. W., & Pope, G. A. 1995. Stochastic Reservoir Modeling Using Simulated Annealing and Genetic Algorithm. *SPE Formation Evaluation*, 4, 7.

Simpson, T. W., Toropov, V., Balabanov, V., & Viana, F. A. 2008. Design and analysis of computer experiments in multidisciplinary design optimization: a review of how far we have come or not. In *12th AIAA/ISSMO multidisciplinary analysis and optimization conference,* 5, 10-12.

Spence, G. H., Couples, G. D., Bevan, T. G., Aguilera, R., Cosgrove, J. W., Daniel, J. M., & Redfern, J. (2014). Advances in the study of naturally fractured hydrocarbon reservoirs: a broad integrated interdisciplinary applied topic. *Geological Society, London, Special Publications*, *374*(1), 1-22.

Stein, M. 1987. Large sample properties of simulations using Latin hypercube sampling. *Technometrics*, *29*(2), 143-151.

Stone, H. L. 1973. Estimation of three-phase relative permeability and residual oil data. *Journal of Canadian Petroleum Technology*, **12**, 53.

Subbey, S., Mike, C., & Sambridge, M. 2003. A strategy for rapid quantification of uncertainty in reservoir performance prediction. Paper SPE 79678 presented at the SPE Reservoir Simulation Symposium, Houston, 3-5 February.

Villadsen, J., Michelson, M. L. 1978. Solution of differential equation models by polynomial approximation. Prentice-Hall.

Witteveen, J. A., Sarkar, S., & Bijl, H. 2007. Modeling physical uncertainties in dynamic stall induced fluid–structure interaction of turbine blades using arbitrary polynomial chaos. *Computers & structures*, *85*(11), 866-878.

Wriedt, J., Deo, M., Han, W. S., & Lepinski, J. 2014. A methodology for quantifying risk and likelihood of failure for carbon dioxide injection into deep saline reservoirs. *International Journal of Greenhouse Gas Control*, *20*, 196-211.

Xiu, D., & Karniadakis, G. 2003. Modeling uncertainty in flow simulations via generalized polynomial chaos. *Journal of computational physics*, *187*(1), 137-167.

Zhang, Y., & Sahinidis, N. V. 2012. Uncertainty quantification in CO_2 sequestration using surrogate models from polynomial chaos expansion. *Industrial & Engineering Chemistry Research*, *52*(9), 3121-3132.

FIGURE CAPTIONS

Fig. 1. Workflow for constructing data-driven surrogates for fractured carbonate reservoirs using multiple experimentally designed simulations.

Fig. 2. Distribution of permeability in the matrix simulation model of a sector of the Amellago Island Outcrop.

Fig. 3. Network of pervasive background fractures with average fracture intensity of (a) 0.05 m^2/m^3, (b), 0.1 m^2/m^3 and (c) 0.2 m^2/m^3.

Fig. 4. Characterization of fracture properties in the Amellago Island Outcrop. (a) Rose diagram showing strike of pervasive regional fractures. (b) Contoured density of fracture poles based on fractures generated for the 3D reservoir model.

Fig. 5. Upscaled fracture permeabilities corresponding to fracture networks with average intensity of (a) 0.05 m^2/m^3, (b), 0.1 m^2/m^3 and (c) 0.2 m^2/m^3. Fracture networks are upscaled to the geocellular grid of the simulation model using the modified Oda method.

Fig. 6. Summary of parameter sensitivities affecting oil recovery and gas utilisation factor (GUF) during CO_2 WAG. Tornado chart shows the difference in the model response when individual parameters are varied between their minimum and maximum values. Full-physics simulations are carried out using the regional discrete fracture network with fracture intensity of 0.1 m^2/m^3. See table 2 for description of symbols.

Fig. 7. Distribution of matrix oil saturation (a) and gas saturation (b) after 8 cycles of immiscible CO_2 WAG injection using an inverted 5-spot well pattern. Geological layer channelling influences recovery efficiency (a), while, buoyancy influences CO_2 migration to the reservoir top (b).

Fig. 8. Profiles of oil recovery (a) and gas utilisation factor (b for experimentally designed simulations used to train and test the surrogate models. Only 50 simulation results are shown to avoid overlapping.

Fig. 9. aPCE surrogate response surfaces for the oil recovery when (a) fault transmissibility, (b) maximum trapped gas saturation, (c) wettability and (d) vertical fracture permeability multiplier are varied along with the horizontal fracture permeability multiplier. 'FT' refers to fault transmissibility. 'S_{gt}' refers to maximum trapped gas saturation. 'KR' refers to the wettability which varies from -1 (oil-wet) to 1 (water-wet). 'Kfz_{mult}t' refers to the vertical fracture permeability multiplier while 'Kfx_{mult}' refers to the horizontal fracture permeability multiplier. Lower GUF is desired for positive recovery economics.

Fig. 10. aPCE surrogate response surfaces for the gas utilization factor when (a) fault transmissibility, (b) maximum trapped gas saturation, (c) wettability and (d) vertical fracture permeability multiplier are varied along with the horizontal fracture permeability multiplier. 'FT' refers to fault transmissibility. 'Sgt' refers to maximum trapped gas saturation. 'KR' refers to the wettability which varies from -1 (oil-wet) to 1 (water-wet). 'Kfz_{mult}' refers to the vertical fracture permeability multiplier while 'Kfx_{mult}' refers to the horizontal fracture permeability multiplier. Lower GUF is desired for positive recovery economics.

Fig. 11. Model comparison of oil recovery and gas utilisation factor (GUF) between full-physics simulations and surrogate models from polynomial regression (a, d), sparse polynomial chaos expansion (b, e) and arbitrary polynomial chaos expansion (c, f). "Actual" refers to results from full-physics IMEX simulations, while, "predicted" refers to results obtained using data-driven surrogates.

Fig. 12. Relative error response surfaces for the oil recovery when the PR surrogate is compared to the aPCE surrogate. Overall error is minimal but notice for all surfaces that the error is lowest at the corners and highest in the centre of the design space because of the deterministic experimental design method.

Fig. 13. Relative error response surfaces for the gas utilisation factor (GUF) when the PR surrogate is compared to the aPCE surrogate. Overall error is minimal but notice for all surfaces that the error is lowest at the corners and highest in the centre of the design space because of the deterministic experimental design method.

Fig. 14. The relative difference in response surfaces when the PR surrogate is compared to the aPCE surrogate for (a) oil recovery and (b) gas utilisation factor (GUF). Further validation sample points have been added using Latin Hypercube sampling to reduce the deterministic sampling bias. Blue dots refer to actual simulation runs for training (dots at the corners) and validation (random dots within the design).

Fig. 15. Cumulative probability distributions for (a) oil recovery and (b) net gas utilization factor generated from 65000 Monte Carlo simulations using the aPCE model. Oil recovery P10, P50 and P90 is 0.31, 0.34 and 0.37 respectively. GUF P10, P50 and P90 is 0.45, 0.53 and 0.60 respectively.

Fig. 16. Genetic algorithm (GA) optimisation process for the fractured carbonate reservoir model. Note the occasional sub-optimal solutions during optimisation to ensure that the GA obtains the optimal global solution. The algorithm is set to maximise the oil recovery, thereby concurrently minimising the GUF.

Fig. 17. Multiple simulation iterations using aPCE surrogate model coupled with genetic algorithm for (a) optimisation of oil recovery and (b) optimisation of net gas utilisation factor.

TABLE CAPTIONS

Table 1. Rock and fluid properties used in reservoir simulation

Table 2. Main parameters used to generate oil-water and gas-oil relative permeability and capillary pressure curves with Corey equations.

Table 3. Parameter, symbols and ranges of the uncertain parameters varied in the experimental design. Matrix relative permeability and capillary pressure curves that indicate the wettability (KR) are represented by discrete variables. '-1' corresponds to oil-wet, '0' corresponds to mixed-wet and '1' corresponds to water-wet.

Table 4. Goodness of Fit Measures. R^2 is the coefficient of determination which indicates how well the data-driven surrogates predict full-physics simulation results. "R^2_{adj}" is a modified form of the coefficient of determination which accounts for the number of regression coefficients in the surrogate equations. RMSE is the root mean square error of the data-driven surrogate compared to the actual simulation.

Table 5. Mean value of uncertain input parameters and outputs (oil recovery factor, RF and gas utilisation factor, GUF) during optimisation with genetic algorithm. Each generation consists of 50 aPCE surrogate evaluations. Optimum solution is obtained after 50 generations.

952

953 # TABLES

954 **Table 1.**

Parameter	Value	Unit
Grid dimension	34 x 35 x 36	-
Grid block size	15 x15 x 3	m
Reservoir pressure	20,684	kPa
Bubble point pressure	11,376	kPa
Oil density	1000	kg/m^3
Water density	800	kg/m^3
Gas density	1.28	kg/m^3
Reservoir temperature	121	^0C

955

956

957 **Table 3.**

Parameter	Symbol	Low	Intermediate	High
Fracture Permeability Multiplier X	Kfx_{mult}	0.1	5.0	10.0
Fracture Permeability Multiplier Y	Kfy_{mult}	0.1	5.0	10.0
Fracture Permeability Multiplier Z	Kfz_{mult}	0.1	5.0	10.0
Fault Transmissibility	FT	0.0	0.5	1.0
Matrix Wettability	KR	-1.0	0.0	1.0
Maximum Trapped Gas Saturation	S_{gt}	0.0	0.2	0.4

958

959

960 **Table 4.**

Goodness of Fit Measure	**Polynomial Regression**		**Sparse Polynomial Chaos**		**Arbitrary Polynomial Chaos**	
	Recovery	GUF	Recovery	GUF	Recovery	GUF
R^2	0.9635	0.9823	0.9768	0.9903	0.9770	0.9903
R^2_{adj}	0.9361	0.9690	0.9594	0.9830	0.9597	0.9830
RMSE	0.0052	0.0098	0.0042	0.0073	0.0042	0.0073

961

962

963

964 **Table 5.**

Generation	Kfx_{mult}	Kfy_{mult}	Kfz_{mult}	FT	KR	S_{gt}	IL_x	IL_y	Inj_{Rate}	RF	GUF
1	1.4057	1.4057	1.4057	0.3059	-1	0.1844	18	17	1736	0.3661	0.4292
2	5.5560	5.5560	5.5560	0.4917	0	0.1373	16	18	1821	0.3867	0.3899
3	1.8103	1.8103	1.8103	0.6586	1	0.1986	17	18	1896	0.4444	0.3593
4	1.6540	1.6540	1.6540	0.5768	1	0.3155	18	20	1814	0.4464	0.3838
5	3.7237	3.7237	3.7237	0.6547	1	0.2877	17	18	1903	0.4467	0.3654
6	1.1319	1.1319	1.1319	0.6648	1	0.1995	17	19	1954	0.4575	0.3537
7	1.0484	1.0484	1.0484	0.7328	1	0.2018	17	18	1957	0.4602	0.3530
8	1.2237	1.2237	1.2237	0.6542	1	0.3243	17	21	1953	0.4723	0.3648
9	1.9328	1.9328	1.9328	0.6916	1	0.3688	15	21	1951	0.4732	0.3620
10	1.2223	1.2223	1.2223	0.8040	1	0.3243	13	21	1956	0.4808	0.3557
11	0.6790	0.6790	0.6790	0.8777	1	0.3348	13	18	1964	0.4797	0.3605
12	0.8006	0.8006	0.8006	0.8771	1	0.3566	14	21	1939	0.4829	0.3603
13	0.6183	0.6183	0.6183	0.8671	1	0.3882	16	21	1978	0.4881	0.3641
14	0.9967	0.9967	0.9967	0.8762	1	0.3892	13	21	1976	0.4899	0.3566
15	0.2848	0.2848	0.2848	0.9702	1	0.4117	14	21	1982	0.4961	0.3599
16	0.2863	0.2863	0.2863	0.9758	1	0.4122	13	21	1983	0.4971	0.3579
17	0.2474	0.2474	0.2474	0.9694	1	0.4125	14	21	1982	0.4969	0.3589
18	0.1059	0.1059	0.1059	0.9992	1	0.4184	13	21	1987	0.4993	0.3581
19	0.1132	0.1132	0.1132	0.9916	1	0.4182	13	21	1986	0.4991	0.3580
20	0.1975	0.1975	0.1975	0.9504	1	0.4117	13	21	1983	0.4978	0.3578
21	0.1047	0.1047	0.1047	0.9979	1	0.4185	13	21	1987	0.4993	0.3581
22	0.1547	0.1547	0.1547	0.9985	1	0.3659	13	21	1987	0.4975	0.3932
23	0.1062	0.1062	0.1062	0.9854	1	0.4112	13	21	1987	0.4993	0.3580
24	0.1021	0.1021	0.1021	0.9991	1	0.4156	13	21	1987	0.4994	0.3581
25	0.1036	0.1036	0.1036	0.9973	1	0.4128	13	21	1987	0.4994	0.3581
26	0.1044	0.1044	0.1044	0.9979	1	0.4097	13	21	1987	0.4994	0.3582
27	0.1034	0.1034	0.1034	0.9978	1	0.4122	13	21	1987	0.4994	0.3581
28	0.1040	0.1040	0.1040	0.9877	1	0.4062	13	21	1987	0.4994	0.3581
29	0.1018	0.1018	0.1018	0.9976	1	0.4079	13	21	1987	0.4994	0.3582
30	0.1014	0.1014	0.1014	0.9968	1	0.4056	13	21	1987	0.4994	0.3582
35	0.1004	0.1004	0.1004	0.9937	1	0.4103	13	21	1987	0.4994	0.3581
40	0.1003	0.1003	0.1003	0.9938	1	0.4098	13	21	1987	0.4994	0.3581
45	0.1002	0.1002	0.1002	0.9898	1	0.4013	13	21	1987	0.4995	0.3581
50	0.1000	0.1000	0.1000	0.9887	1	0.4014	13	21	1987	0.4995	0.3581

965

966

967

ALGORITHM

Algorithm 1: **Genetic algorithm for optimisation by selection, mutation and recombination**

Start
  Initialize solutions $\mathbf{x}_i$ of population $\boldsymbol{\lambda}$
  Evaluate objective function for the solutions $\mathbf{x}_i$ in $\boldsymbol{\lambda}$
  Repeat
    For $i = 0$ to $\boldsymbol{\beta}$
      Select $\boldsymbol{\rho}$ parents from $\boldsymbol{\lambda}$
      Create new $\mathbf{x}_i$ by recombination
      Mutate $\mathbf{x}_i$
      Evaluate objective function for $\mathbf{x}_i$
      Add $\mathbf{x}_i$ to $\boldsymbol{\lambda}$'
    Next
    Select $\boldsymbol{\mu}$ parents from $\boldsymbol{\lambda}$' and form new $\boldsymbol{\lambda}$
  Until termination condition
End

FIGURES

Figure 1

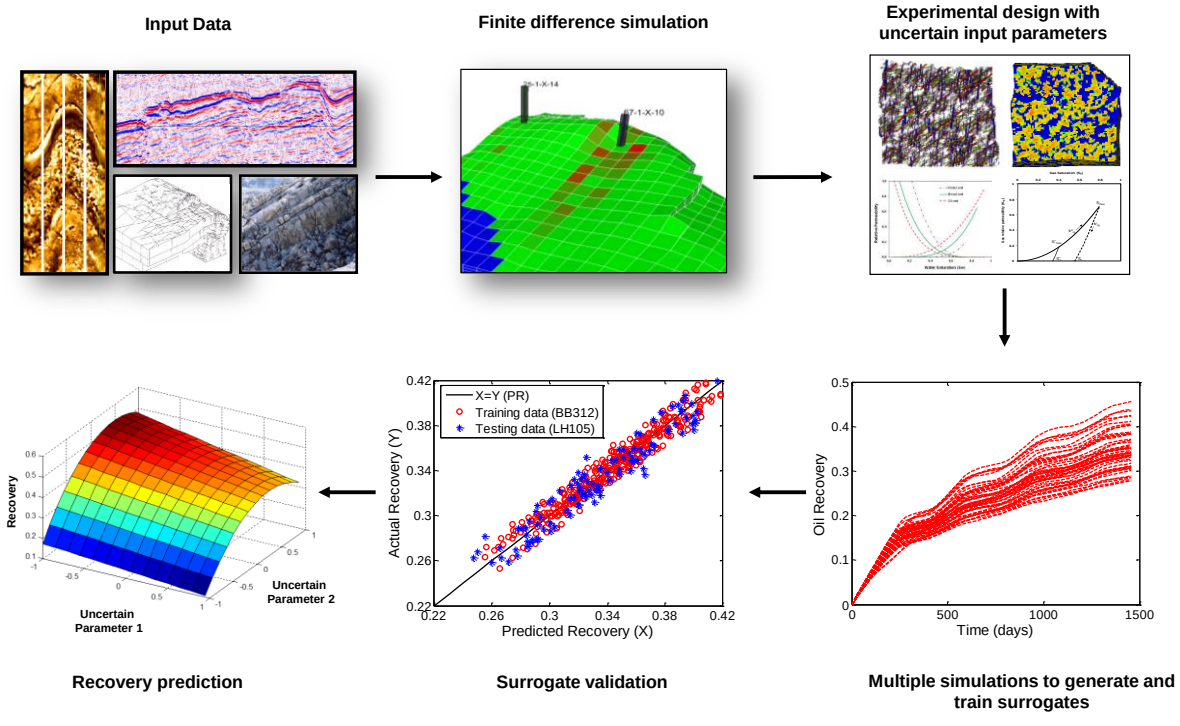


Figure 2.

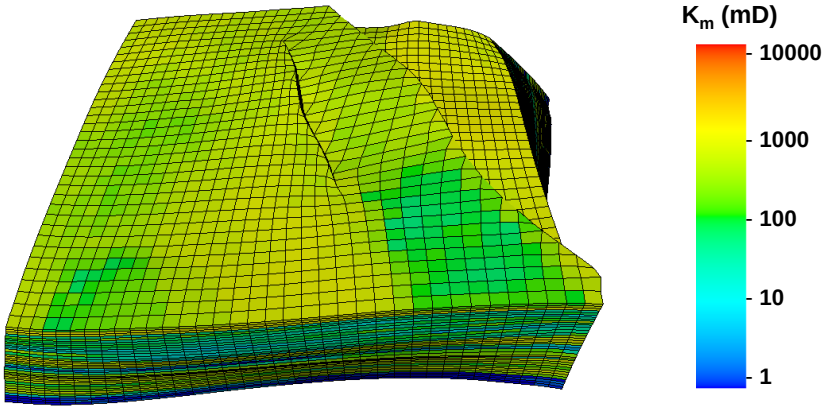


Figure 3.

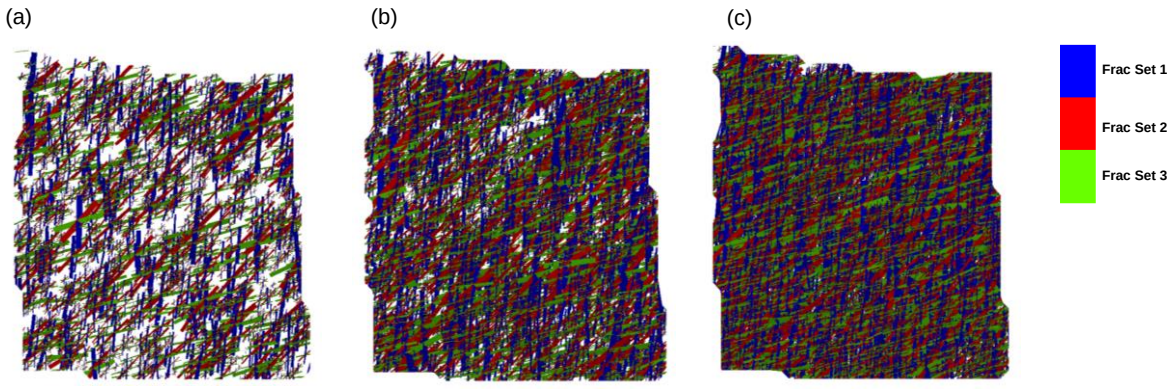


Figure 4.

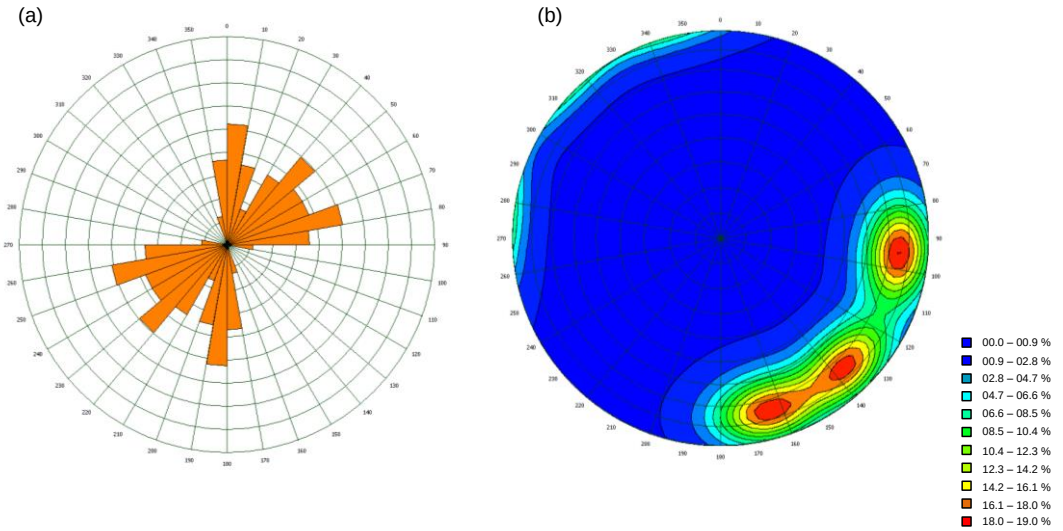


Figure 5.

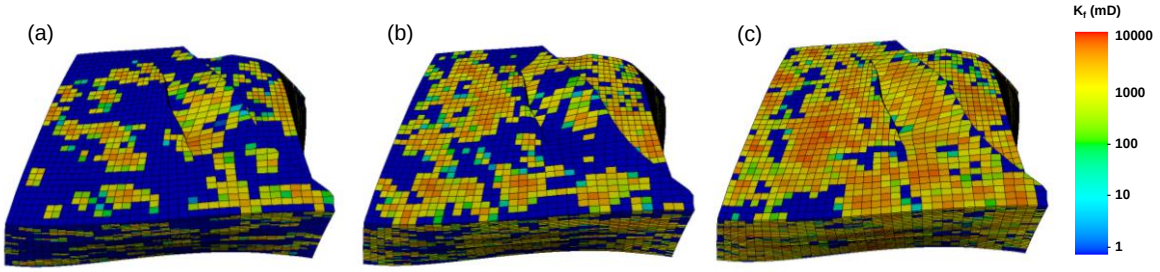


Figure 6.

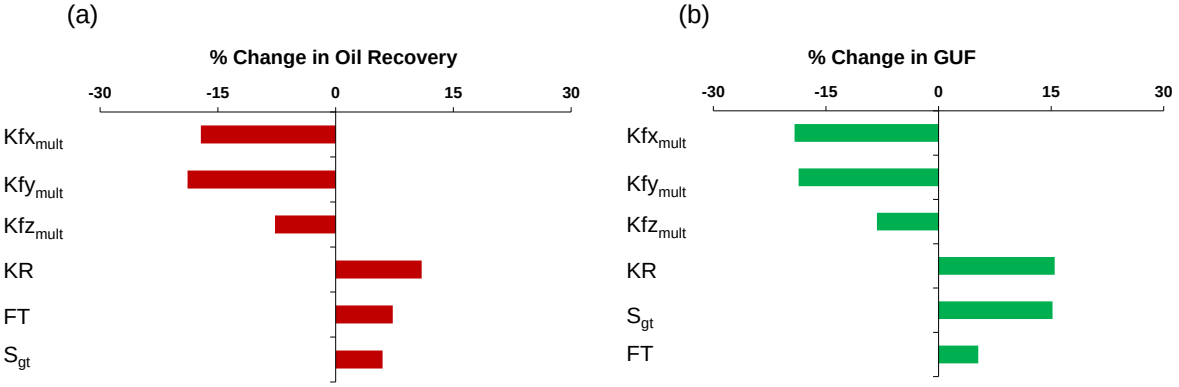


Figure 7.

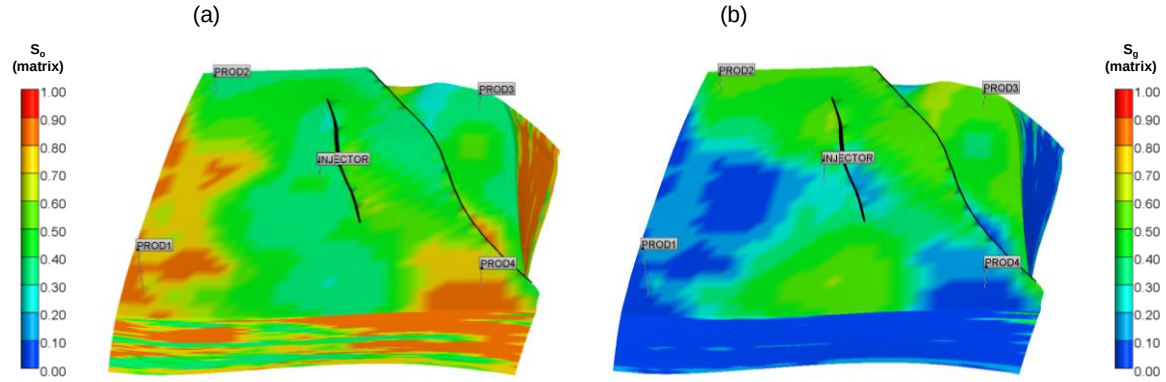


Figure 8.

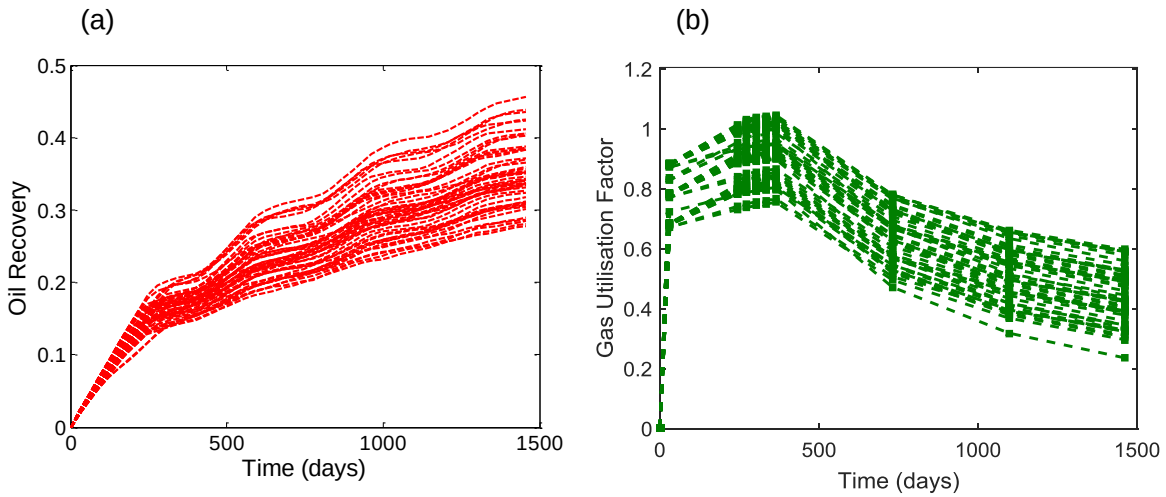


Figure 9.

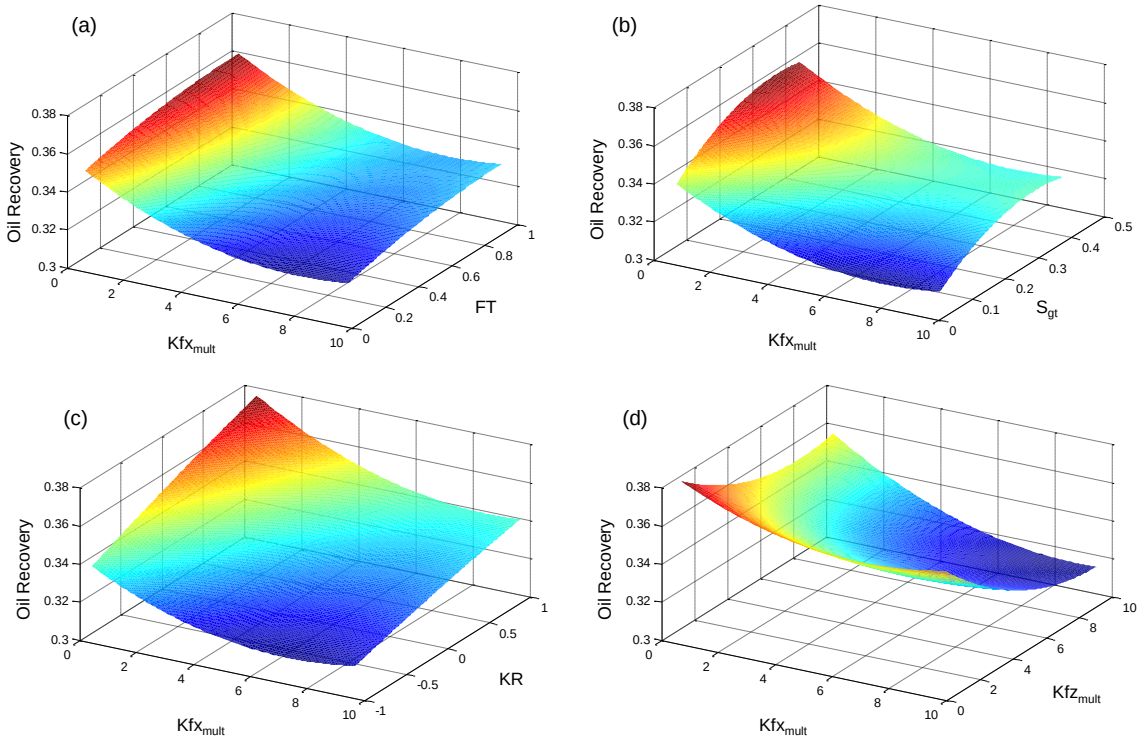


Figure 10.

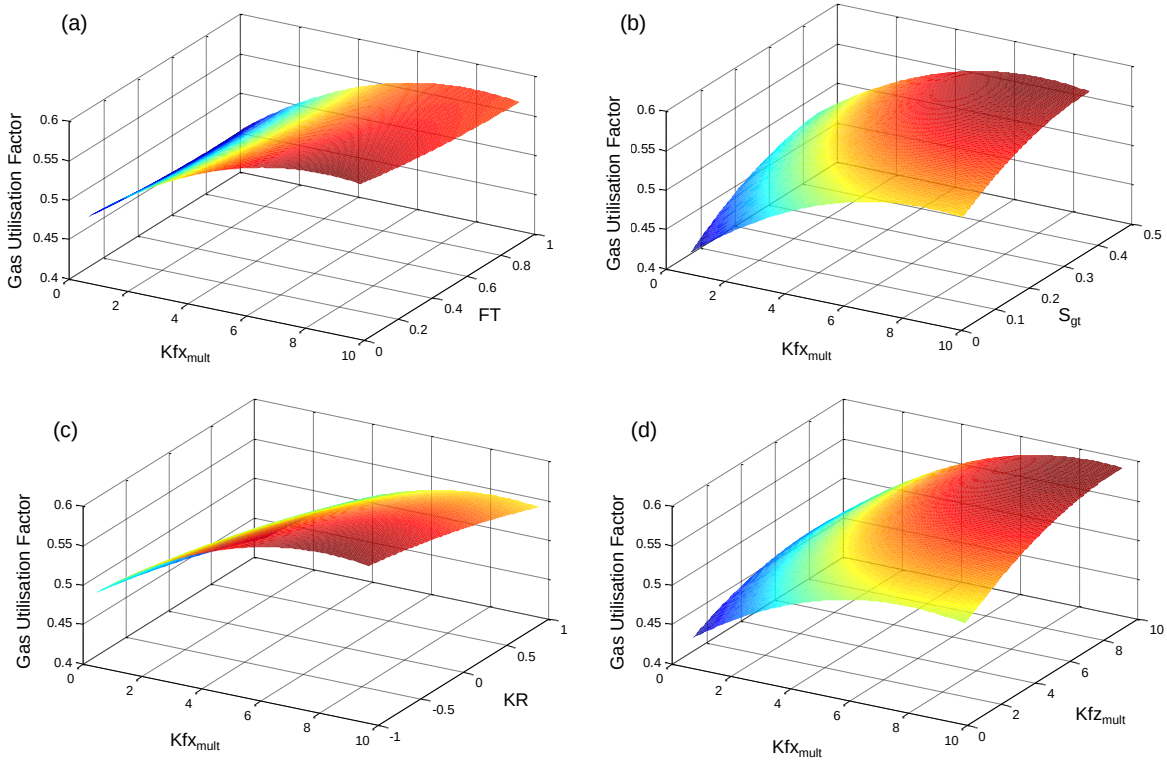


Figure 11.

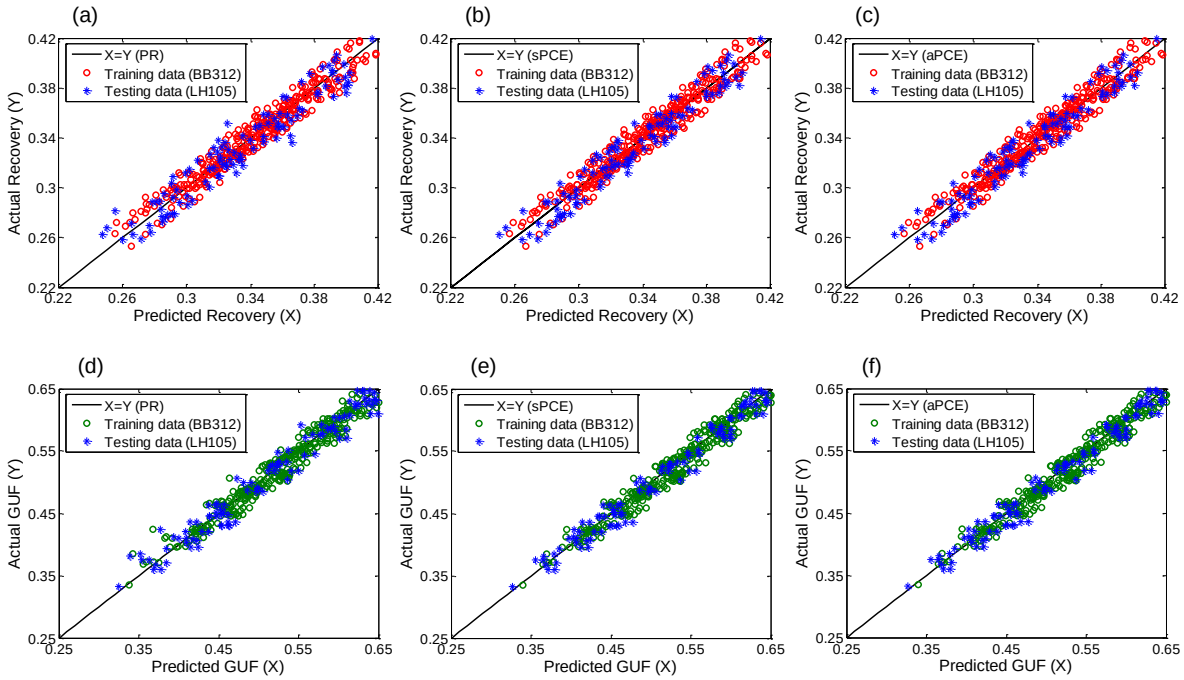


Figure 12.

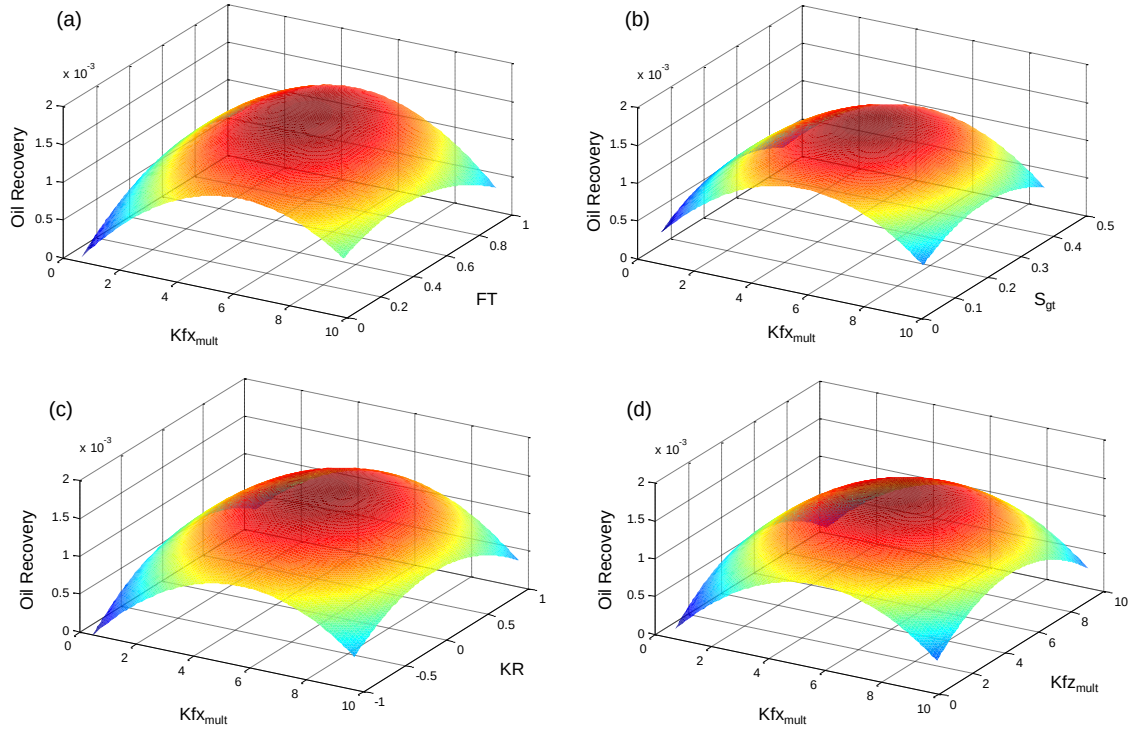


Figure 13.

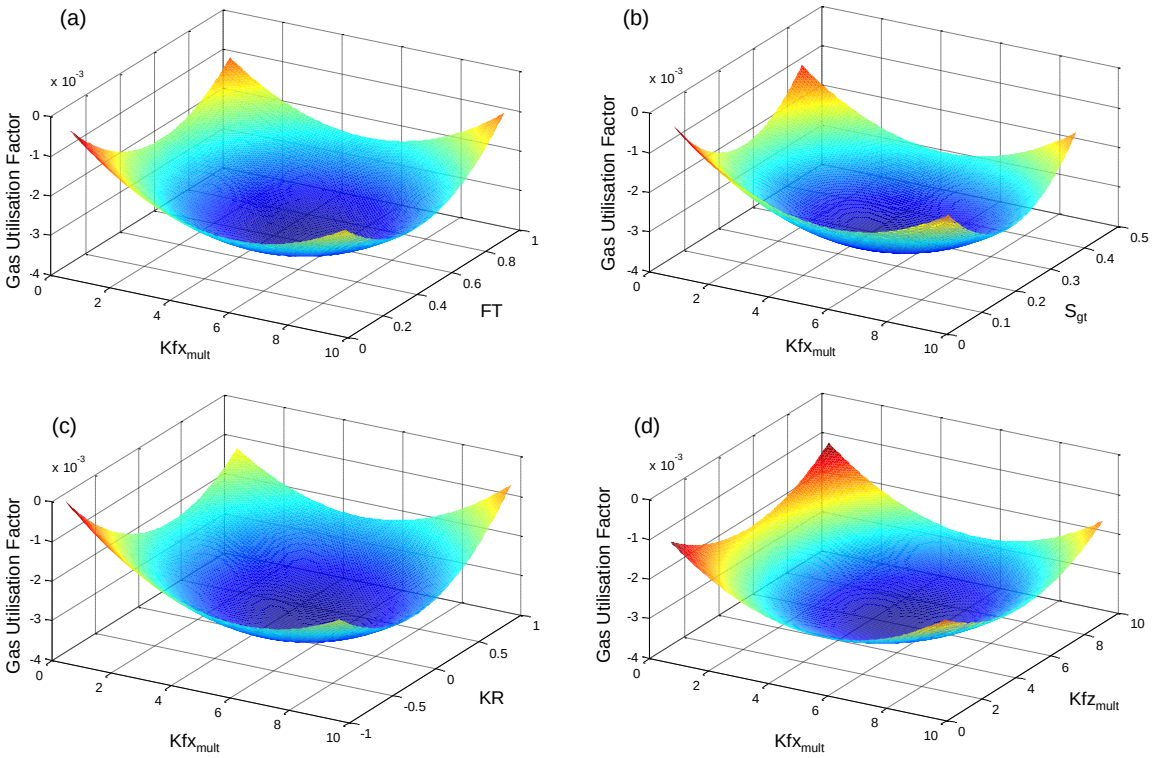


Figure 14.

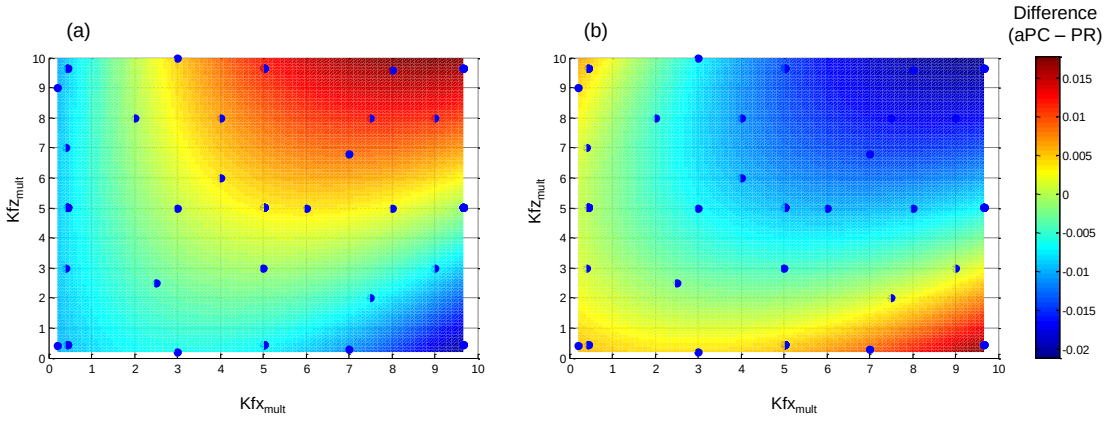


Figure 15.

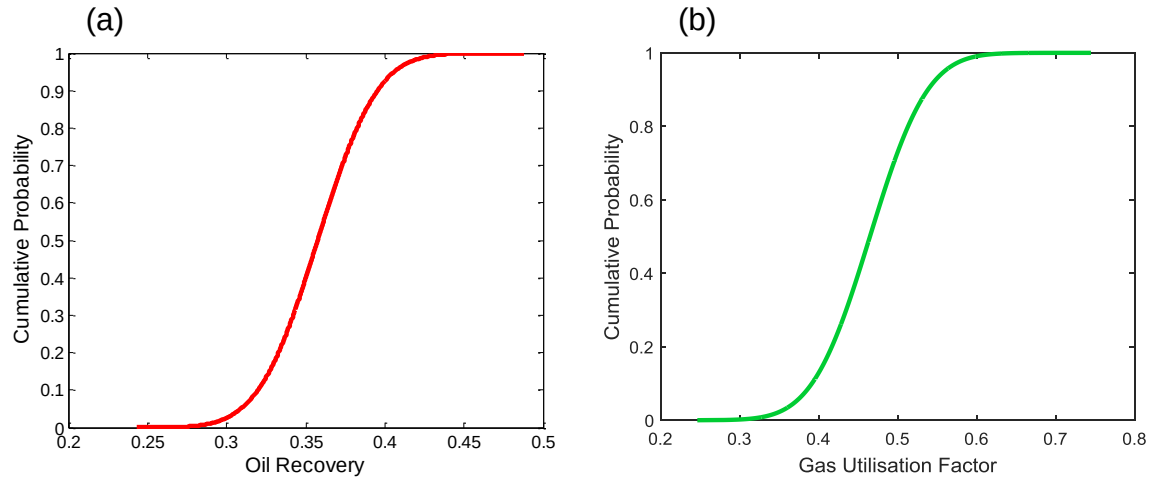


Figure 16.

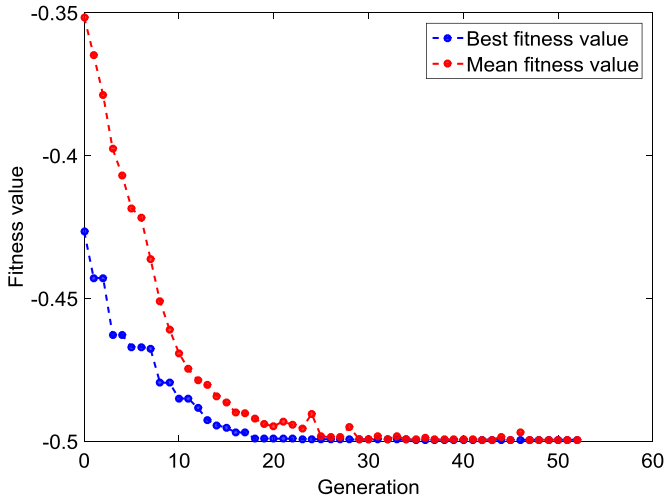
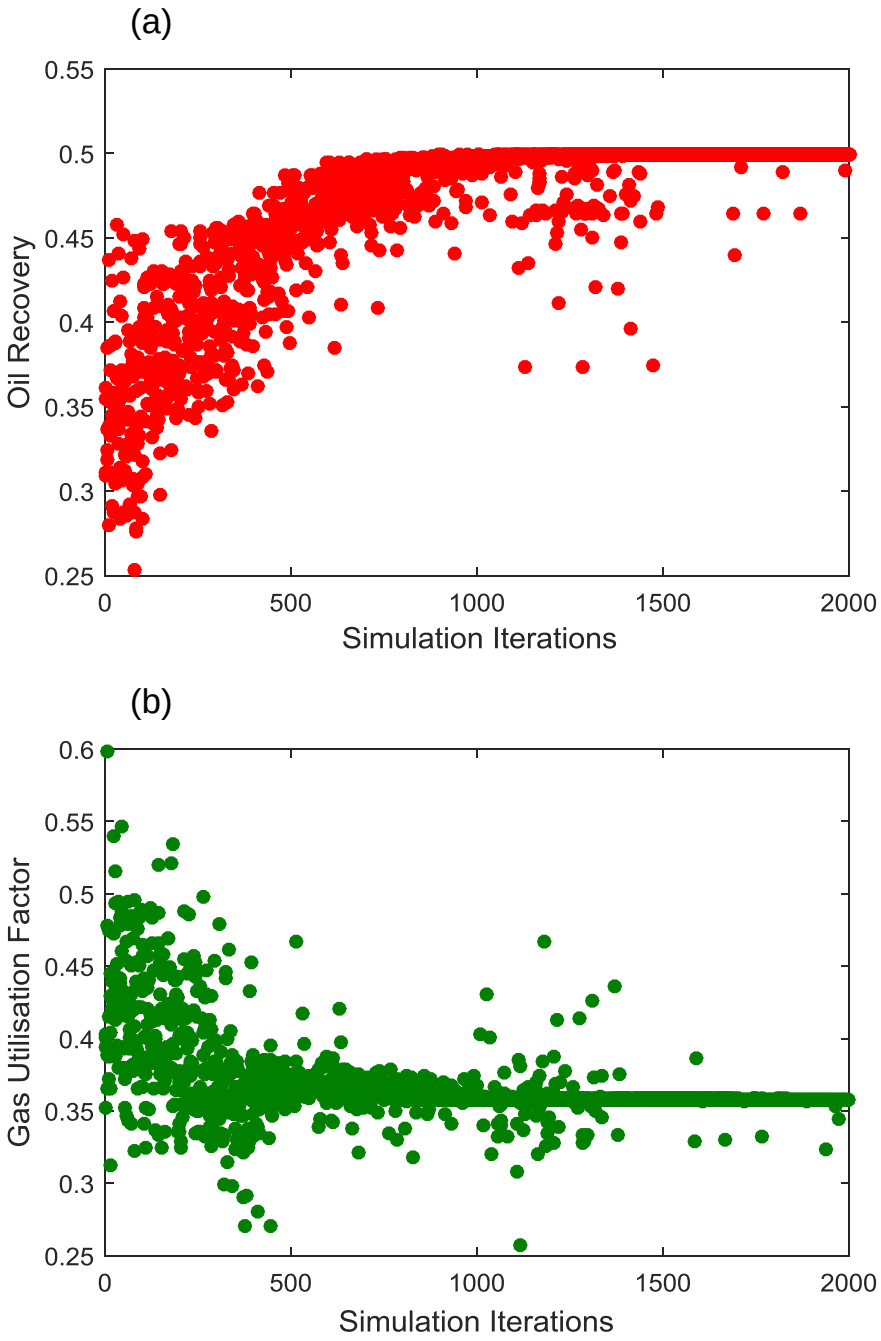


Figure 17.


(a)
Oil Recovery
0.55
0.5
0.45
0.4
0.35
0.3
0.25
0
500
1000
1500
2000
Simulation Iterations
(b)
Gas Utilisation Factor
0.6
0.55
0.5
0.45
0.4
0.35
0.3
0.25
0
500
1000
1500
2000
Simulation Iterations